DOCUMENT RESUME

ED 386 475 TM 024 029

AUTHOR	Zwick, Rebecca
TITLE	An Analysis of Graduate School Careers in Three Universities: Differences in Attainment Patterns across Academic Programs and Demographic Groups. GRE Board Professional Report No. 86-21P.
INSTITUTION	Educational Testing Service, Princeton, NJ. Graduate Record Examination Board Program.
SPONS AGENCY	Graduate Record Examinations Board, Princeton, N.J.
REPORT NO	ETS-RR-91-1
PUB DATE	May 91
NOTE	125p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Doctoral Programs; *Educational Attainment; *Foreign Students; *Graduate Students; Higher Education; *Minority Groups; Sex Differences
IDENTIFIERS	Graduate Record Examinations

ABSTRACT

The graduate careers of nearly 5,000 Ph.D.-seeking students from 11 departments in each of 3 major universities were investigated with a special focus on minority students. Minorities and women were found to be underrepresented in graduate school, and to have generally lower candidacy and graduation rates than their white and male counterparts. In two of the three schools, foreign students had higher candidacy and graduate rates than did white Americans. Also, in two of the three schools, the percentage of foreign students increased substantially in recent years. A more general finding was that the candidacy and graduation rates in the 8 years following matriculation were higher in quantitatively-oriented departments than in the humanities and social sciences. In general, undergraduate grades and Graduate Record Examinations scores had only a minimal association with the attainment of candidacy and graduation. Among these academically select students, nonacademic factors may play a crucial role in determining who ultimately attains the doctoral degree. Eleven tables and 42 figures present study data. Five appendixes contain the study cover letter, a project description, a description of data requirements, 17 supplementary tables, and an empirical Bayes survival analysis. (Contains 42 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 386 475

GRE®
RESEARCH

An Analysis of Graduate School Careers in Three Universities: Differences in Attainment Patterns Across Academic Programs and Demographic Groups

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Rebecca Zwick

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
H. I. BRAUN
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

May 1991

GRE Board Professional Report No. 86-21P
ETS Research Report 91-17

BEST COPY AVAILABLE

Educational Testing Service, Princeton, New Jersey

An Analysis of Graduate School Careers in Three Universities: Differences in Attainment Patterns Across Academic Programs and Demographic Groups

Rebecca Zwick

GRE Board Report No. 86-21P

May 1991

This report presents the findings of a research project funded by and carried out under the auspices of Educational Testing Service and the Graduate Record Examinations Board.

Educational Testing Service, Princeton, N.J. 08541

The Graduate Record Examinations Board and Educational Testing Service are dedicated to the principle of equal opportunity, and their programs, services, and employment policies are guided by that principle.

Graduate Record Examinations and Educational Testing Service are U.S. registered trademarks of Educational Testing Service; GRE, ETS, and the ETS logo design are registered in the U.S.A. and in many other countries.

Copyright © 1991 by Educational Testing Service. All rights reserved.

Contents

Acknowledgments v

Abstract vii

Introduction 1

Data Collection And Editing 3

Data Analyses 7

Summary and Discussion 73

Appendix A Letter of Request for Participation 77

Appendix B Description of Project 81

Appendix C Description of Data Requirements 85

Appendix D Supplementary Tables 89

Appendix E Empirical Bayes Survival Analysis 109

References 117

Acknowledgments

This project is part of the Minority Graduate Education research agenda, jointly sponsored by Educational Testing Service (ETS) and the Graduate Record Examinations (GRE) Board. I am grateful for their sponsorship. I also wish to express my appreciation to the administration and staff of the participating schools. In addition, I would like to thank Ka-Ling Chan, Thomas Florek, Jo-Ling Liang, Dorothy Thayer, and Minhwei Wang for statistical programming and data analysis; Kate Pashley and Ira Sample for preparing the figures that appear in the report; Henry Braun for providing consultation on the survival analyses and reviewing the report; and Foster Schoenthaler for sharing his expertise on the GRE data base. Finally, I would like to thank Peg Goertz, Charlotte Kuh, Toni Clewell, and the GRE Research Committee for their helpful comments and Debbie Kline for her expert editorial assistance.

Abstract

The graduate careers of nearly 5,000 Ph.D.-seeking students from 11 departments in each of three major universities were investigated, with a special focus on minority students. Minorities and women were found to be underrepresented in graduate school and to have generally lower candidacy and graduation rates than their White and male counterparts. In two of the three schools, foreign students had higher candidacy and graduation rates than did White Americans. Also, in two of the three schools, the percentage of foreign students increased substantially in recent years.

A more general finding was that the candidacy and graduation rates in the eight years following matriculation were higher in quantitatively oriented departments than in the humanities and social sciences. In general, undergraduate grades and Graduate Record Examinations scores had only a minimal association with the attainment of candidacy and graduation. Among these academically select students, nonacademic factors may play a crucial role in determining who ultimately attains the doctoral degree.

Introduction

In recent years, the status of graduate education in the United States has received much attention in the popular press, as well as in educational journals. Studies show that American students represent a decreasing percentage of students earning doctorates in U.S. graduate schools, particularly in science and engineering (ACE, 1987; National Research Council, 1986; Thomas, 1987; Trent & Copeland, 1987). Another disturbing trend is the decrease in the participation of Black American students in graduate education during the last decade (ACE, 1987; ACE/ECS, 1988; Blackwell, 1987; Brown, 1987; Mooney, 1989; Trent & Copeland, 1987). Some educators have expressed concern that talented undergraduates may be choosing to go to professional schools or may be entering the work force immediately upon graduation (Brademas, 1984; Hartnett, 1987) and that this may be particularly true for minority students (Mooney, 1989; Pruitt & Isaac, 1985). The underrepresentation of women in graduate education continues to be a major concern as well (Chamberlain, 1988). These phenomena have major implications for the future role of minorities and women in university faculties and in the American work force in general.

To effect changes in the current system of graduate education, it is necessary to understand what happens to individuals who do enroll in graduate school. At what pace do these students reach milestones in their graduate careers, such as advancement to Ph.D. candidacy and completion of the doctoral degree? Can standard preadmissions measures, such as test scores and undergraduate grades, be used to distinguish graduate students who complete the doctorate from those who do not? How do the patterns of achievement differ across academic programs, universities, and demographic groups? Answers to these questions can be useful to graduate school administrators in allocating resources and improving educational policies that affect admissions and retention.

Using data on graduate students at Northwestern University, Zwick and Braun (1988) investigated these questions, focusing on 14 departments: Chemical Engineering, Computer Science, Chemistry, Mathematics, Physics, Counseling Psychology, Clinical Psychology, Sociology, Theatre, English, History, Political Science, Economics, and Philosophy. Only students who stated at entry that they were seeking a Ph.D. were included in the study.

Among the most prominent findings was the increase in the percentage of foreign students enrolled in these 14 departments at Northwestern—from 15% in 1975-1977 to 32% in 1984-1986. Increases were most noteworthy in the Computer Science and Physics departments. The patterns of attainment of graduate school milestones, such as Ph.D. candidacy and graduation, were examined for each department, as well as for gender and ethnic groups. There was substantial variation across departments and, to a lesser degree, across demographic groups. The highest candidacy and graduation rates occurred in Clinical Psychology and Chemistry, the lowest in Theatre and Computer Science. In general, graduation rates for foreign students were higher than those for

U.S. citizens. The association between the attainment of milestones and measures of academic potential, such as undergraduate grade-point average (UGPA) and Graduate Record Examination (GRE) scores was also investigated. The likelihood of attaining candidacy or of completing a doctorate was found to bear little relation to UGPA and GRE scores. This finding is probably a result of the use of UGPA and GRE in the selection of students into graduate departments. Evidently, within this highly able group of students, these conventional measures of academic skills could not distinguish between those who did and did not achieve candidacy and graduation.

The present study extends the work of Zwick and Braun (1988) in two ways. First, the new study includes data from three institutions. This aspect of the research serves to provide some information about the generalizability of the Zwick and Braun findings and about the feasibility of multi-institution research on graduate education. The second way in which the present study differs from the previous one is that it includes more extensive analyses of graduate school careers for Blacks, Whites, and foreign students and for men and women, and, unlike the Zwick and Braun study, includes some analyses based on Asian and Hispanic students.

The analyses for this project were designed to meet two principal objectives:

1. One goal of the present study was to determine how the patterns of attainment of Ph.D. candidacy and the Ph.D. degree differ across departments and across institutions. How large is the variability across institutions in patterns of degree attainment compared to the variability among departments within a single school? To what degree is the attainment of candidacy and of the Ph.D. degree predicted by Graduate Record Examinations (GRE) scores and undergraduate grade-point average (UGPA)? Does the association of candidacy attainment and degree completion with these preadmissions measures differ across departments and institutions?

2. Another major purpose of the study was to explore in detail the graduate careers of Asian, Black, and Hispanic students who are United States citizens. As a first step, an examination was made of the demographic composition of Ph.D. seekers who entered any of 11 departments at the three institutions during the years 1978 through 1985. Then, an analysis of the patterns of candidacy and degree attainment for minority students was conducted in selected clusters of departments. The attainment patterns of these students were compared to those for White and foreign students. The achievement of candidacy and the Ph.D. degree was also examined separately for men and women. Finally, the association of preadmissions measures of academic skills with candidacy and graduation was examined within demographic groups.

Data Collection And Editing

School Contacts

Between March 1987 and December 1987, 20 universities were contacted to determine whether they maintained data bases that would lend themselves to the planned analyses and whether they would be willing to participate in the study. In most cases, a letter (Appendix A), accompanied by a brief description of the project (Appendix B), was first sent to the school, followed by a telephone call to the graduate division. Schools that expressed interest in the project were sent a more detailed description of the data requirements (Appendix C). The primary requirement was that certain key information be available on the individual student level for at least five consecutive years of enrollment. Schools were considered for inclusion if most of the following data were available: entry date, department, citizenship, ethnicity (at least for U.S. citizens), gender, UGPA, GRE scores, date of advancement to candidacy, and date of graduation. Furthermore, it was required that data be in machine-readable form and that the data base be centralized (i.e., that it be available through a single university office). Schools were offered reimbursement for costs associated with the preparation and mailing of data tapes.

Six of the 20 schools that were contacted initially agreed to participate. In most cases, refusals were due to the unavailability of the required data. It was subsequently determined that two of the original six participating schools could not supply data on the variables of interest. A third school withdrew because no staff were available to create the needed data tapes. This left three participating schools, all of which are large research universities. The schools requested that they not be identified in research reports and are therefore referred to only as Schools 1, 2, and 3.

Selection of Students for Study

In determining which entry cohorts and departments to include in the analyses, the primary consideration was that comparable data be available for all three schools. Data from Ph.D.-seeking students who entered during the period beginning in the fall of 1978 and ending in the fall of 1985 were included in the analyses; these were the entry cohorts for which all three schools provided data.

It was important to select departments that were similarly defined in all three schools and that had adequate numbers of students for analysis. These criteria led to the choice of the following 11 departments: Chemistry, English, History, Mathematics, Political Science, Psychology, Economics, Philosophy, Physics, Computer Science, and Sociology. In Phase 2 of the analyses, which focused on ethnic and gender groups, departments were grouped for analyses, as described in a later section.

Linking of School 1 Records to GRE Files to Obtain GRE Scores

For the years prior to 1982, the data base from School 1 did not include GRE scores. Therefore, with the university's permission, records from School 1 were linked to the GRE data base at ETS to obtain the scores for students in the departments selected for study. The linking process also provided the opportunity to acquire some information about undergraduate grades for students at School 1, which does not include UGPA in its student records: Students who take the GRE are asked to respond to a series of background items, including a question about their undergraduate grades in the last two years of college. The responses, which are in terms of letter grades, were converted to numerical values for use in this study. This UGPA information differs from that obtained from the other two schools in that it corresponds to only the last two years of college and was provided by the students themselves in the context of a testing situation, rather than obtained from school transcripts.

The GRE data base also served as a source of ethnicity data for students who were missing this information in the School 1 data base. Of the 2,913 School 1 students in this study, the GRE data base was the source of GRE scores for 1,479, UGPA data for 2,073, and ethnicity data for 127.

Coding of Ethnicity

Encoding student ethnic status in a manner that was consistent across schools was important in this study. The mapping of the information provided by the three schools and the GRE data base into the seven categories used in this study is detailed in Table 1. The need for consistency and for adequate sample sizes necessitated the use of a classification system that was, in some cases, less refined than those used by the participating schools. For example, it was not possible to conduct separate analyses for Puerto Rican and Chicano students or for Chinese and Japanese students. The ethnic categories used in this study were Asian, Black, Hispanic, and White. These codes were applied to U.S. citizens only; foreign students constituted a separate category. Students who were neither Asian, Black, Hispanic, White, nor foreign were given a code of "other"; students for whom information was unavailable were placed in a "missing" category. As described above, missing ethnic information for School 1 students was in some instances filled in from the GRE data base.

Data Screening

Special computer programs were developed for editing the data provided by the schools. These programs implemented checks for duplicate records, inconsistent information on dates (such as candidacy dates before entry dates

Table 1

Scheme for Encoding Ethnic and Citizenship Data[a]

Analysis Category	Categories from Schools and GRE Data Base: School 1	GRE Data Base[b]	School 2	School 3
Asian	Chinese Japanese Korean Polynesian Vietnamese/Other Asian	Oriental/Asian American	Oriental/Asian American	Oriental
Black	Black	Black/Afro-American	Black/Afro-American	Black
Hispanic	Chicano Other Spanish	Mexican-American/ Chicano Puerto Rican Other Hispanic or Latin American	Mexican-American/ Chicano Puerto Rican (Mainland) Puerto Rican (Commonwealth)	Hispanic Mexican-American/ Chicano Puerto Rican
White	White	White	White/Caucasian	White
Other	American Indian East Indian/Pakistani Filipino Other	American Indian/Eskimo Aleut Other	American Indian/Alaskan Native Other	American Indian
Foreign	[c]	[c]	Foreign	[c]
Missing	Decline to state [Blank]	[Blank]	Decline to state [Blank]	[Blank]

[a] The ethnic descriptions in the body of the table are those used by the data sources. Note that the GRE program no longer uses the designation "Oriental."

[b] Information from the GRE data base was used for some School 1 students who were missing ethnic data.

[c] Students were classified as foreign based on a citizenship variable.

or after Ph.D. dates), and out-of-range values, such as impossible GRE scores.[1] A total of nine records were deleted because of major data errors or inconsistencies. The final data base for the project included 4,637 student records: 2,913 from School 1, 633 from School 2, and 1,091 from School 3.

[1]GRE scores from 1981 or earlier that were greater than 800 were recoded to 800. The 1987-1988 *GRE Guide* (ETS, 1987) notes that "when comparing verbal and quantitative scores earned and reported after October, 1981, with verbal and quantitative scores earned earlier, any earlier scores higher than 800 should be interpreted as being equivalent to 800" (p. 29).

Data Analyses

Overview

The data analyses conducted for this study were of three basic types. First, descriptive analyses were conducted, showing the numbers of students entering each of the 11 departments and the proportions of women, minority group members, and foreign students. The second category of analyses involved the examination of patterns of attainment of graduate school milestones for each department and for demographic groups within clusters of departments. The final type of analysis involved investigation of the association between attainment of milestones and potential explanatory variables.

Descriptive Analyses

As a first step in analyzing the data, the demographic makeup of the 11 selected departments was investigated. Results were tabulated separately for two cohorts of students, each corresponding to roughly half of the matriculation period included in this study. The first cohort consisted of students who entered during the period beginning in the fall of 1978 and ending in the summer of 1981; the second cohort consisted of students with entry dates during the period beginning in the fall of 1981 and ending in the fall of 1985.

Table 2 provides information about the percentages of ethnic and gender groups for the three schools and the two cohorts, combined across the 11 departments. Tables D-1 through D-11 (Appendix D) provide corresponding information separately for each of the 11 departments.

Ethnicity Results

In terms of ethnicity, the three schools (all departments combined; see Table 2) showed similar patterns for Cohort 1: 72 to 75% of the students were White and 18 to 20% were foreign. The percentages of students who were Asian Americans, Black Americans, and Hispanic Americans summed to only 5 to 7%.[2] In Cohort 2, the School 1 results remained virtually unchanged, while at School 2 and School 3, the percentage of Whites decreased by roughly 10 and the percentage of foreign students increased by about the same amount.

The ethnic group results for individual departments (Tables D-1 through D-11) showed that the most typical pattern of change was a decrease in the

---

[2]The figures in Tables 2 and D-1 through D-11 have been rounded to the nearest percent. The discussion here is based on a more precise set of results.

Table 2

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	1058	1855	210	423	423	668
Ethnic Group						
Asian	3	4	1	2	1	2
Black	2	1	2	0	3	2
Hispanic	2	3	1	1	1	1
White	72	71	75	63	73	64
Other	2	1	0	0	0	0
Foreign	18	19	19	31	20	31
Missing	1	1	1	2	1	0
Gender Group						
Male	74	72	72	65	72	73
Female	26	28	28	35	28	27

[a] Percentages may not add to 100 because of rounding.

[b] Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

percentage of White students and a corresponding increase in the percentage of foreign students. This pattern occurred in most departments at School 2 and School 3 and was also evident in the Computer Science and Sociology departments at School 1. Significant exceptions[3] to this pattern occurred in the Political Science and Philosophy departments at School 1 and the English, History, and Sociology departments at School 3. In these departments, the percentage of White students increased and the percentage of foreign or Black students decreased.

The scarcity of Asian, Black, and Hispanic American students is one of the most striking aspects of the results. Only rarely did any one of these groups comprise more than 5% of the students in a particular department within a cohort. The percentage of Asian Americans exceeded 5 at School 1 in the Chemistry and English departments in Cohort 1 and in the Psychology, Philosophy, and Computer Science departments in Cohort 2 and in the School 2 Political Science department in Cohort 1 (one student). The percentage of Black students exceeded 5 in the School 1 Sociology department in Cohort 2 and the School 3 Sociology department in Cohort 1; in the School 2 Political Science department in Cohort 1 and the School 3 Political Science department in Cohorts 1 and 2; in the School 2 History department in Cohort 1 (one student) and the School 3 History department in Cohorts 1 and 2; in the School 3 English department in Cohort 1; and in the School 1 Psychology department in Cohort 1. The percentage of Hispanic Americans was greater than 5 in Cohorts 1 and 2 in the School 1 Sociology department, in Cohort 1 in the School 3 Sociology department, and in Cohort 1 in the School 2 Mathematics department (one student).

Gender Results

The proportions of men and women in graduate school are also worthy of examination. Based on the gender results for the combined departments (Table 2), the percentages of male students were remarkably consistent, ranging from 72 to 74%, except for Cohort 2 at School 2, where the percentage of men dropped to 65%. A separate analysis (not shown) revealed that, combined across cohorts, the percentage of males at each of the three schools was highest for foreign students, followed in order by Asians, Whites, Hispanics, and Blacks.

Considering the data from all three schools, the most heavily male departments were Mathematics, Physics, and Computer Science, all of which were typically at least 80% male. The departments that came closest to having equal numbers of men and women were English, History, Psychology, and Sociology. The most notable changes in the gender composition of departments were the increases in the percentages of women in the Philosophy department at School 1, the Chemistry and Psychology departments at School 2, and the Computer Science department at School 3, and the decreases in the percentages

[3]In very small departments, shifts in demographic composition that are small in absolute terms correspond to large changes in percentages. Such changes are not noted in the discussion.

of women in English and Sociology at School 1 and in History at School 3.

Patterns of Attainment of Graduate School Milestones

For purposes of monitoring academic programs and for making projections about the U.S. work force, it is useful to examine rates of graduation for students in different universities, academic programs, and demographic groups. Investigating the attainment of candidacy can also be informative: How likely is it that students will attain this first milestone?

One complicating factor in analyses of this kind is that requirements for candidacy and graduation differ across universities and, in some cases, across academic programs within universities. In addition to differences in documented policies, there may be less formalized differences in the ways in which schools or departments implement certain policies. For example, in some departments, it may be standard procedure to require all students in a given entry year to take qualifying examinations at the same time. Those who pass, therefore, have the same candidacy date. In contrast, other departments may have extremely flexible schedules for taking exams.

Some information on the differences among the policies of the three schools in this study can be obtained from school catalogs. At School 1, admission to candidacy requires successful completion of qualifying examinations, including an oral component. Students must apply for advancement to candidacy no later than the semester after the one in which the qualifying examination was passed. The period during which students are allowed to be "in candidacy" is limited by regulations that vary across departments. A minimum of two years or four semesters in residence is required for the Ph.D. degree.

At School 2, doctoral students must pass a comprehensive examination not later than the session prior to the session of graduation. Based on consultation with School 2 graduate division staff, the date of completion of this exam was considered to be the candidacy date for purposes of this study. Doctoral students are expected to have completed at least three years of residence in a graduate college. Students must take an oral final exam after submitting their dissertations and must pass the exam no later than five years after passing the comprehensive exam.

At School 3, admission to Ph.D. candidacy is contingent on completion of departmental requirements, including a comprehensive qualifying examination. Students are expected to be admitted to candidacy before the end of the third calendar year after initial registration and must be admitted to candidacy by the end of the twelfth quarter after registration. All requirements for the doctoral degree must be met within 5 years of admission to candidacy or within 8 years of the last year of consecutive full-time residency or within 10 years of initial registration, whichever comes first. Students may petition for a two-year extension of the deadline.

Variations in policy across schools and departments undoubtedly had an impact on the analyses conducted for this project. It is important, therefore, to regard the patterns of attainment of graduate school milestones

as functions of academic policies, as well as student characteristics.

Survival Analysis

A detailed picture of the rates of attainment of candidacy and graduation at multiple time points can be achieved through survival analysis, a statistical method used to model the time until the occurrence of some event. Although survival analysis has its origins in medical research, where the events of interest are typically the deaths of individuals, the method has recently gained popularity in other fields. Two examples from the field of education are an analysis of Ph.D. attainment at Stanford University (Mathematical Methods in Educational Research, 1983) and an analysis of teachers' career patterns in Michigan public schools (Murnane, Singer, & Willett, 1988).

The goal of survival analysis is estimation of the survival function, S(x), which is the probability that an event will take more than x units of time to occur. In this study, the events of interest are the graduate school milestones, candidacy and graduation, and the units of time are the number of years since entry into graduate school. For example, suppose that S(5) is estimated to be .60 for a survival analysis of graduation. This means that if a student has completed five years of graduate school, the estimated probability that he or she has not yet graduated is .60.

The survival function, S(x), is defined as follows:

$$S(x) = P(X > x) = 1 - F(x),$$

where X is the random variable representing the time elapsed until the milestone is reached and F(x) is the cumulative distribution function of the random variable X, evaluated at X = x.

A related function is the hazard function, h(x), which is the instantaneous risk of the occurrence of an event at X = x, given that the event has not occurred before time x. The hazard function is defined as

$$h(x) = f(x) / S(x),$$

where f(x) is the probability density function. If the hazard function takes on a high value at time x, the survival function will show a correspondingly large drop at time x. In fact, the hazard and survival functions are equivalent ways of summarizing the distribution of survival times, since

$$h(x) = - d/dx [\ln (S(x))].$$

If X is assumed to have the exponential distribution with parameter θ, h(x) is constant and equal to θ.

It is important to note that the statistical terms "survival," "hazard," and "risk" are used here in a way that differs from everyday parlance. In this report, survival refers to the probability of remaining in graduate

school without achieving the event of interest; for example, the probability that the degree is not received by a particular time. Similarly, we speak of the "hazards" or "risks" of attaining candidacy or completing a degree. (Some may find this usage to be counterintuitive in the present context; others may find it appropriate!)

Standard methods exist for both nonparametric and parametric estimation of survival functions (e.g., see Kalbfleisch & Prentice, 1980). Difficulties in estimation can occur when sample sizes are small, however, particularly when it is of interest to estimate separate curves for subpopulations. Bayesian methods can be particularly useful in this situation. By using information from multiple subpopulations, the estimates of individual survival curves for those subpopulations can be stabilized. Whereas previous Bayesian efforts had focused on the estimation of a single survival curve, Braun (1985) developed an empirical Bayes approach for estimating a family of survival functions. A brief description of this approach, along with a summary of the models used in this study, is given in Appendix E. Details of the Bayesian approach and the associated estimation procedures are provided in Braun (1985) and Braun and Zwick (1989); a general description of empirical Bayes methods is given in Braun (1988).

A phenomenon that had to be considered in analyzing these data is censoring: the removal of individuals from the risk set (the group of individuals who are available to experience the event of interest) for reasons other than the occurrence of the event. In this study, some individuals were censored because the data collection effort ended during their graduate careers. Censoring dates for candidacy and graduation in the three schools are given in Table 3. In the survival model applied here, censoring is accommodated through adjustment of the risk set. This means that if the termination of the data collection effort occurs at time x of a student's graduate career, that student will no longer be considered "exposed" or "at risk" for candidacy or graduation after time x. Because all three schools had earlier censoring dates for graduation than for candidacy (see Table 3), candidacy and graduation analyses were, in effect, based on slightly different groups of students. Therefore, the reported results cannot provide accurate estimates of the proportions of students who attained candidacy but did not graduate as of a certain time point. As a result of the differential censoring, it is even possible for estimated graduation rates to exceed estimated candidacy rates. This occurred for the three Chemistry departments at year 8 (see Tables 4 and 5).

Note that, for purposes of the analyses in this study, students who left graduate school without attaining the milestone in question were still considered to be part of the risk set. Roughly speaking, these analyses focused on the probabilities of achieving milestones x years after entering school. Accurate information about student drop-out was not available from the university data bases. If it had been possible to obtain this information, the students who left school without attaining milestones could have been deleted from the risk set. This type of analysis, however, would have had a different interpretation. It would have involved estimation of the

Table 3

Censoring Dates for Survival Analysis

	Last Candidacy Date for Students Included in Study	Last Graduation Date for Students Included in Study
School 1	July 28, 1988	May 20, 1988
School 2	Fall 1987	Summer 1987
School 3	May 30, 1987	August 29, 1986

probabilities of attaining milestones x years after entering school for those students still in school at that time. In an analysis of this kind, the attainment of milestones would have appeared more likely.

Two major categories of survival analyses were conducted. The first category involved comparisons of attainment patterns for the 11 departments in the three schools. The second major category examined attainment patterns for ethnic and gender groups. Within each of these analysis categories, survival analyses were conducted for two events: the achievement of candidacy and the completion of the Ph.D. degree. For each survival analysis, graphs of the empirical Bayes estimates of the survival functions are provided (Figures 1-42). The vertical axis of these graphs represents the estimated percentage of students who have not graduated (100 $\hat{S}(x)$); the horizontal axis represents the number of years since entering school. To facilitate interpretation, survival analysis results are also summarized in Tables 4-9 in terms of the estimated percentage of students who have attained candidacy or graduated as of certain time points (i.e., $100(1 - \hat{S}(x))$).

A preliminary analysis was conducted to determine whether the survival analyses should be conducted separately for the two entry cohorts. Within each department, the means, variances, and intercorrelations of UGPA, GRE verbal score (GREV), GRE quantitative score (GREQ), and a variable indicating the achievement of candidacy were compared across cohorts.[4] Except for some cases in which sample sizes were small, means and variances were very similar across cohorts, and correlations showed no systematic differences across cohorts. Therefore, in order to achieve greater stability of estimation, students from both cohorts were combined for the survival analyses.

Analyses Based on the 11 Departments

The results of the survival analyses for the 11 selected departments are given in Figures 1-18. To facilitate interpretation of the survival analysis graphs, the 11 selected departments have been grouped for display purposes. Group I consists of four quantitatively oriented departments: Chemistry, Physics, Mathematics, and Computer Science. Group II consists of three humanities departments: English, Philosophy, and History. Group III includes four social science departments: Psychology, Political Science, Sociology, and Economics. Sample sizes for the analyses are given in Tables D-12 and D-13. The initial sample sizes for the 11 departments in the three schools ranged from 22 to 648.

---

[4]The candidacy variable is defined in the section below, "Relation of Candidacy and Graduation to Measures of Academic Potential." A graduation indicator variable was not included in the cohort comparison because many members of the later cohort did not have adequate opportunity to graduate before the end of the data collection. Survival analyses of graduation for the later cohort are still meaningful, however, since individuals are removed from the risk set when the end of the data collection occurs, as explained above.

Candidacy Analysis. Figures 1-9 show the survival functions for the candidacy analyses. Figures 1-3 contain graphs for each of the three groups of departments for School 1, Figures 4-6 for School 2, and Figures 7-9 for School 3. As explained above, the survival function at time x is the probability that the event has not occurred by time x. As an aid in interpreting the survival analysis results, Table 4 shows the estimated percentages of students achieving Ph.D. candidacy five and eight years after entry for the 11 departments in the three schools. The tabled quantities are $100(1 - \hat{S}(5))$ and $100(1 - \hat{S}(8))$, where $\hat{S}(x)$ represents the estimated survival function at time x.[5] Five years after entry, School 1 and School 3 students were nearly always more likely to have achieved candidacy than School 2 students; results at eight years show that the School 1 students were, in general, most likely to have achieved candidacy, followed in order by School 3 and School 2 students. Survival curves tended to level off by year 5 at School 2 and by year 4 at School 3, indicating that if candidacy was not reached by these years, it was unlikely to be attained. Only at School 1 were at least 50% of students in all 11 departments estimated to have achieved candidacy by year 8. For several departments, survival functions were still decreasing at year 8.

Group I departments were somewhat more diverse than the other two groups of departments, but within this group, Chemistry generally showed the highest candidacy rates, while Computer Science usually showed the lowest. In all three schools, the hazard functions (not shown) peaked in year 2 for Chemistry and in year 3 for Mathematics, indicating that these were the most likely candidacy years for these departments. For Physics and Computer Science, the hazard functions peaked during years 3 through 5. Peaks in the hazard functions for candidacy correspond to decreases in the survival functions in Figures 1-9.

In Group II, History consistently showed the highest candidacy rates, followed in order by English and Philosophy. At School 3, the hazards for all departments in Group II peaked in year 3; at School 1, they peaked in years 4 and 5, and at School 2, they peaked in years 3 through 5.

In Group III, Psychology had the highest candidacy rates at both five and eight years in all three schools. Political Science had the second highest rate at all three schools at five years; Sociology consistently had the second highest rate at eight years. At School 1, the hazard functions peaked in years 4 through 6. At Schools 2 and 3, the peaks of the hazard functions occurred in year 3, except for Economics at School 2 (year 5).

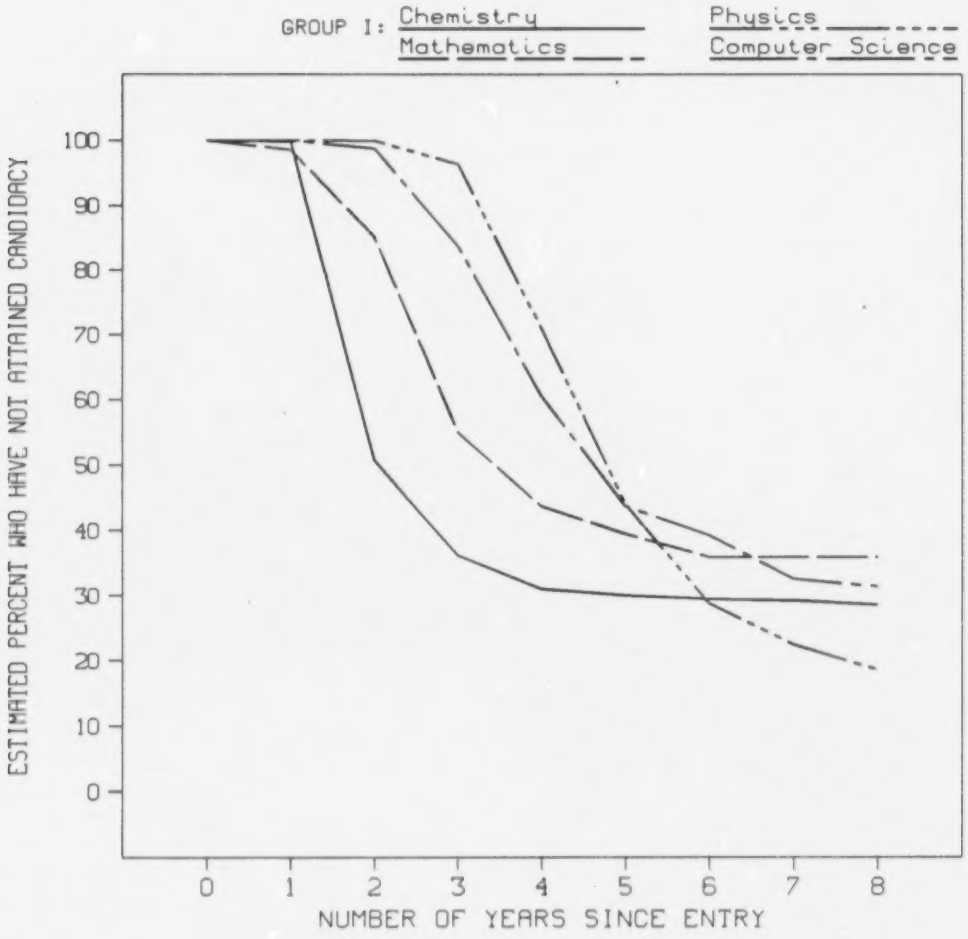
In general, Group I departments had the highest candidacy rates, followed in order by Groups III and II.

---

[5]Five-year values were chosen for tabulation because, in many of the survival analyses, differences across student groups first became apparent at this point. Eight-year values were tabled because this is the latest time point for which survival estimates could be obtained.

FIGURE 1

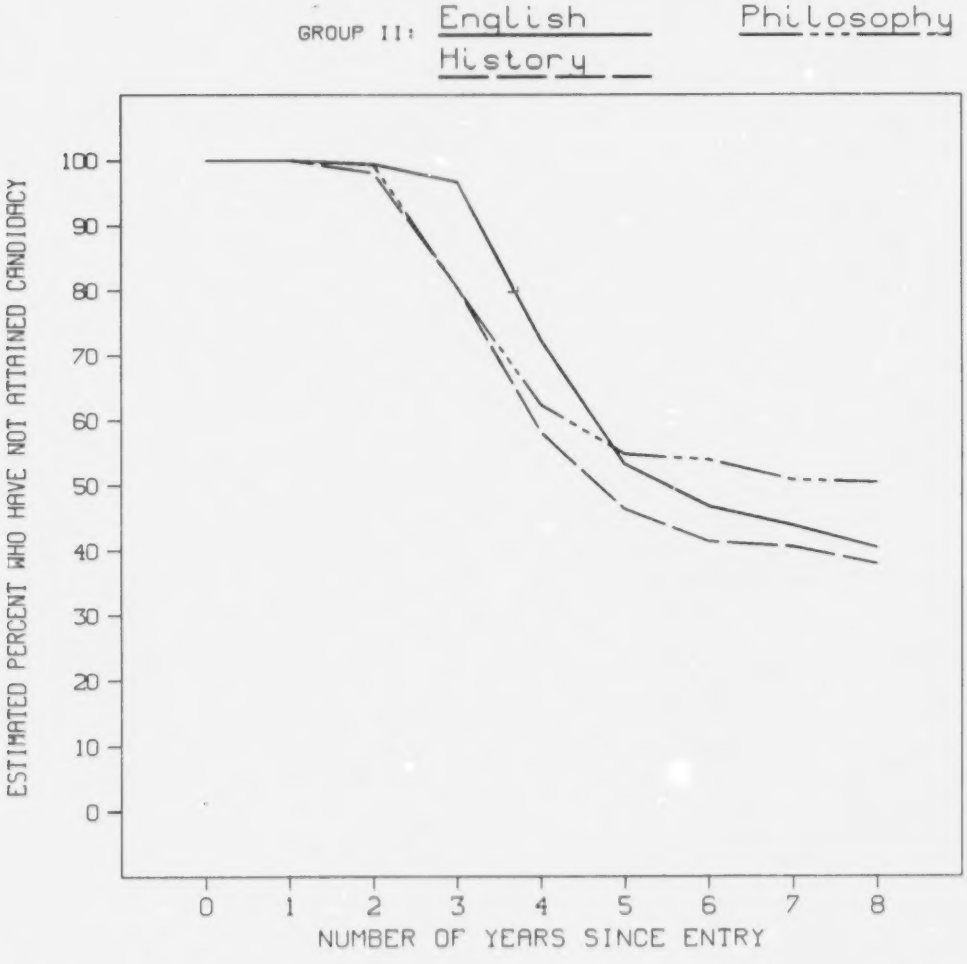
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 2

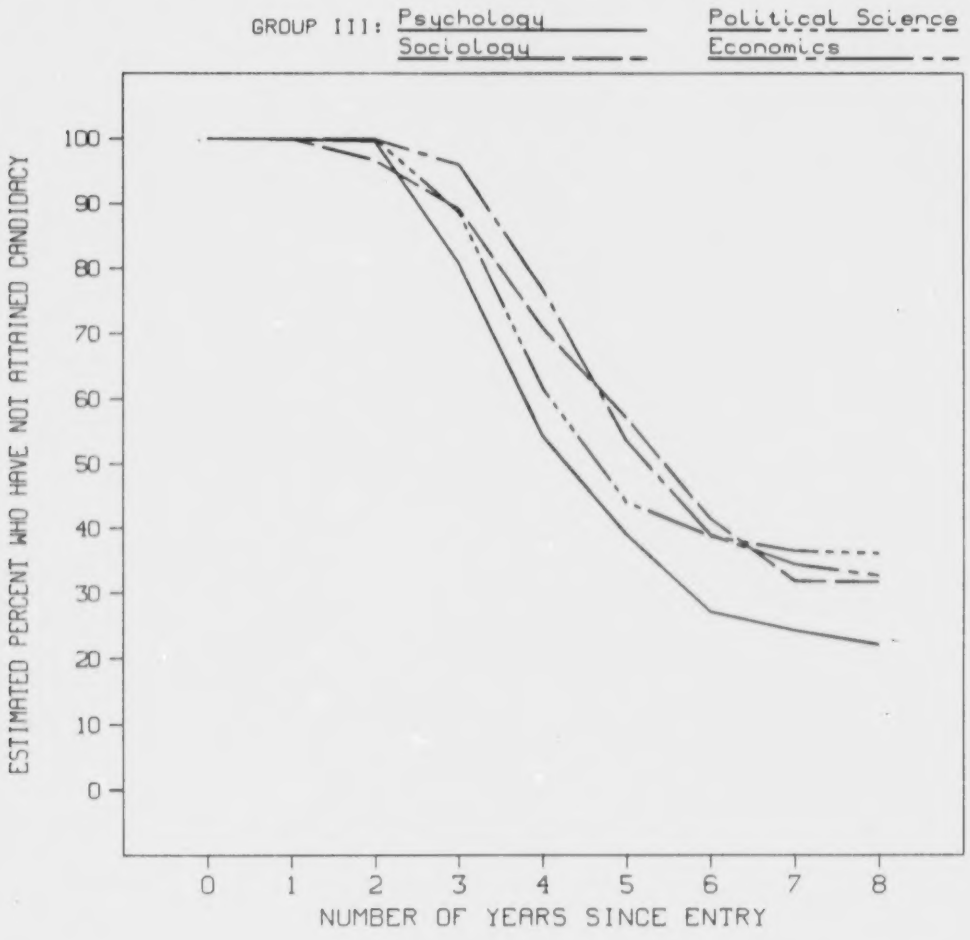
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 3

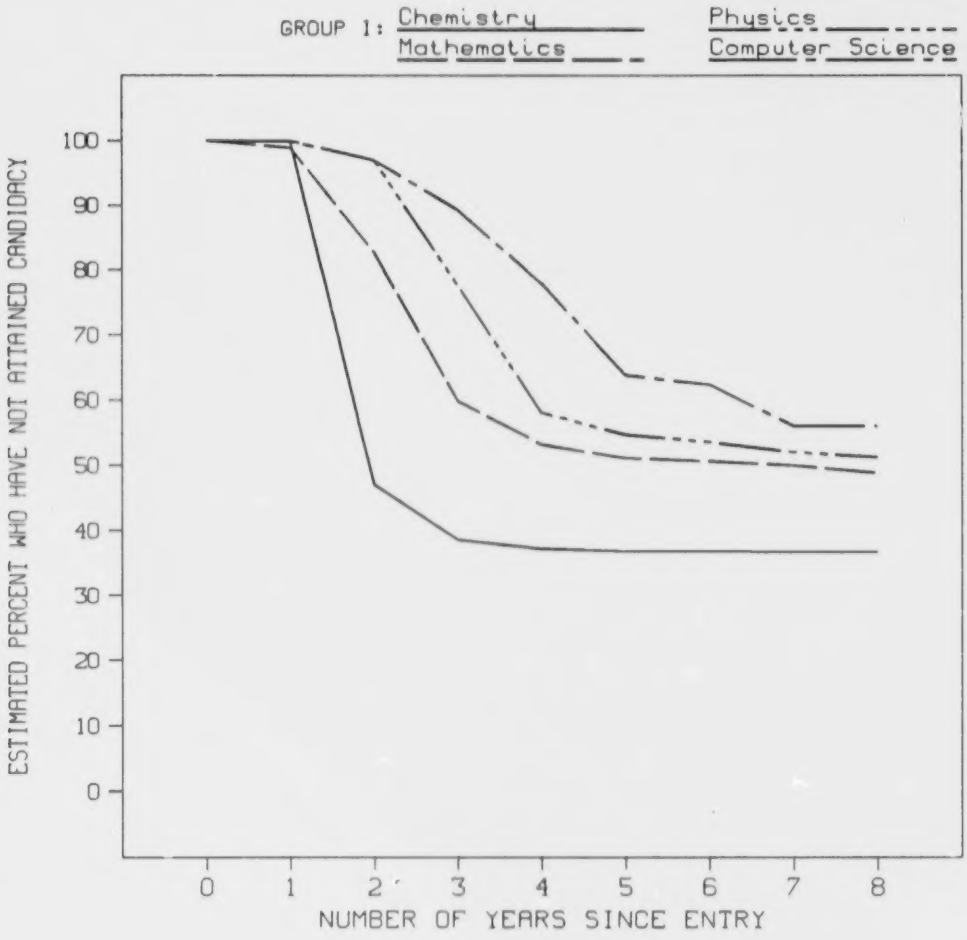
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 4

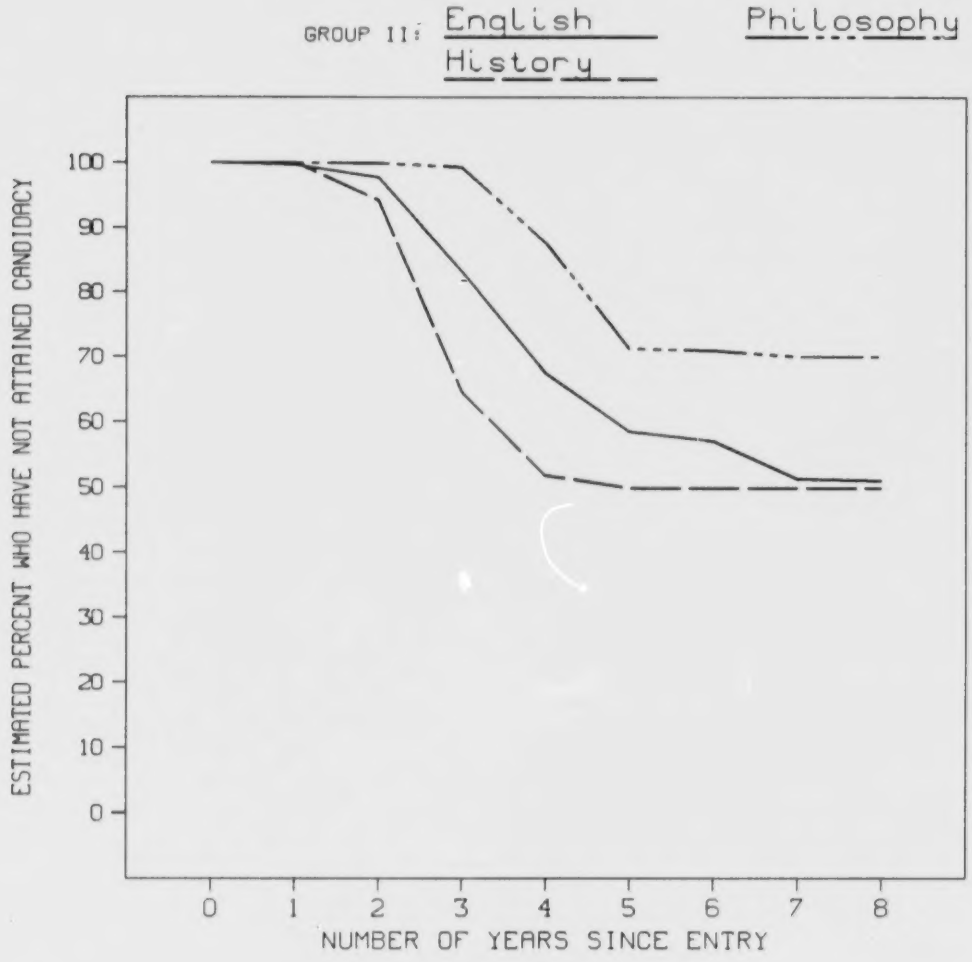
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 5

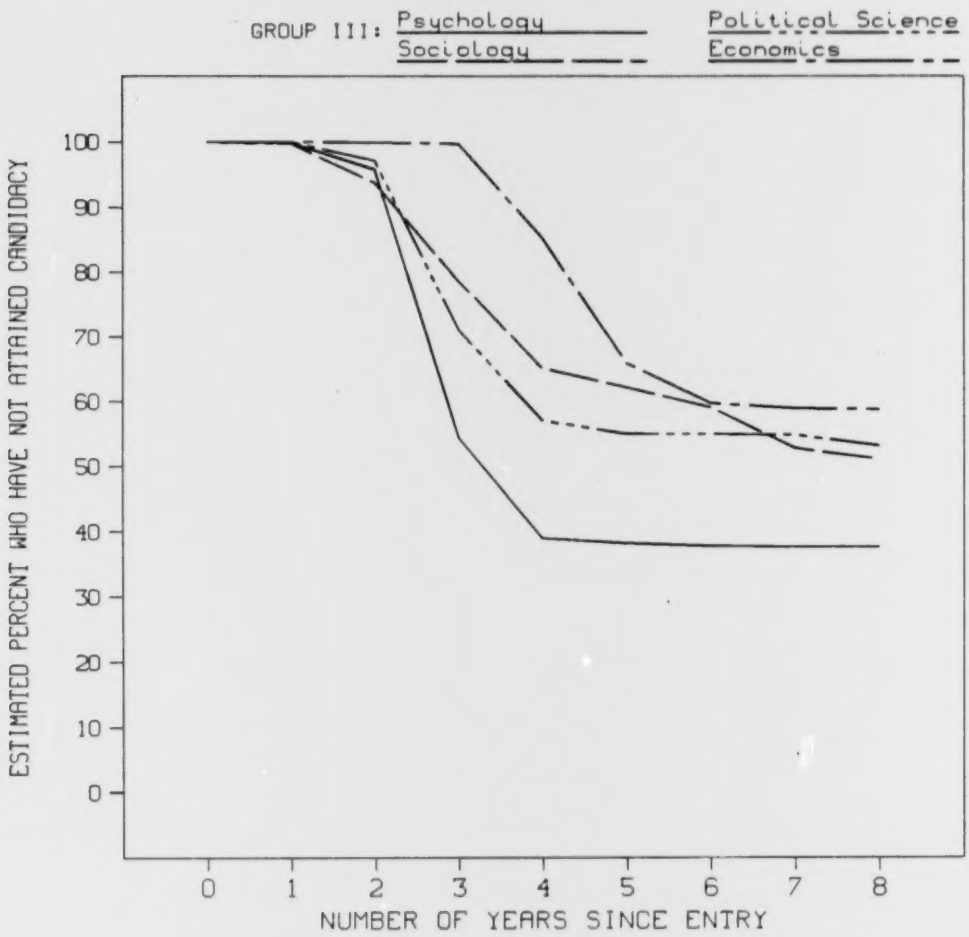
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 6

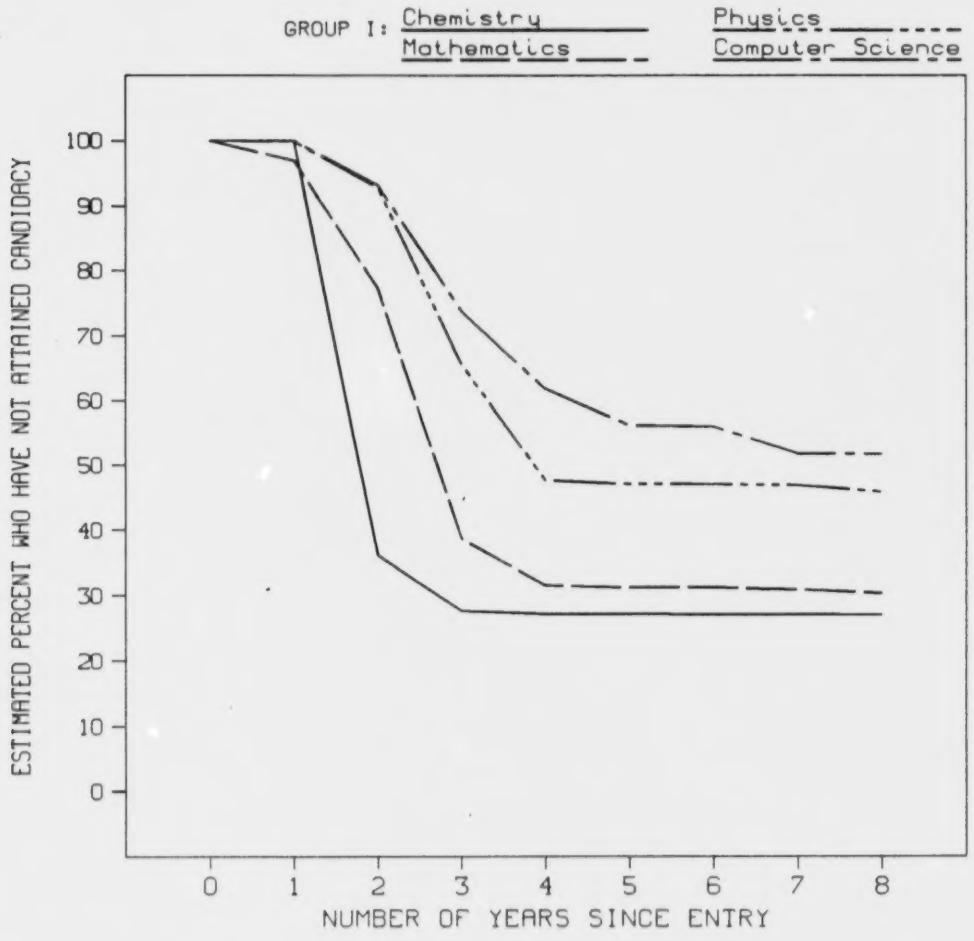
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 7

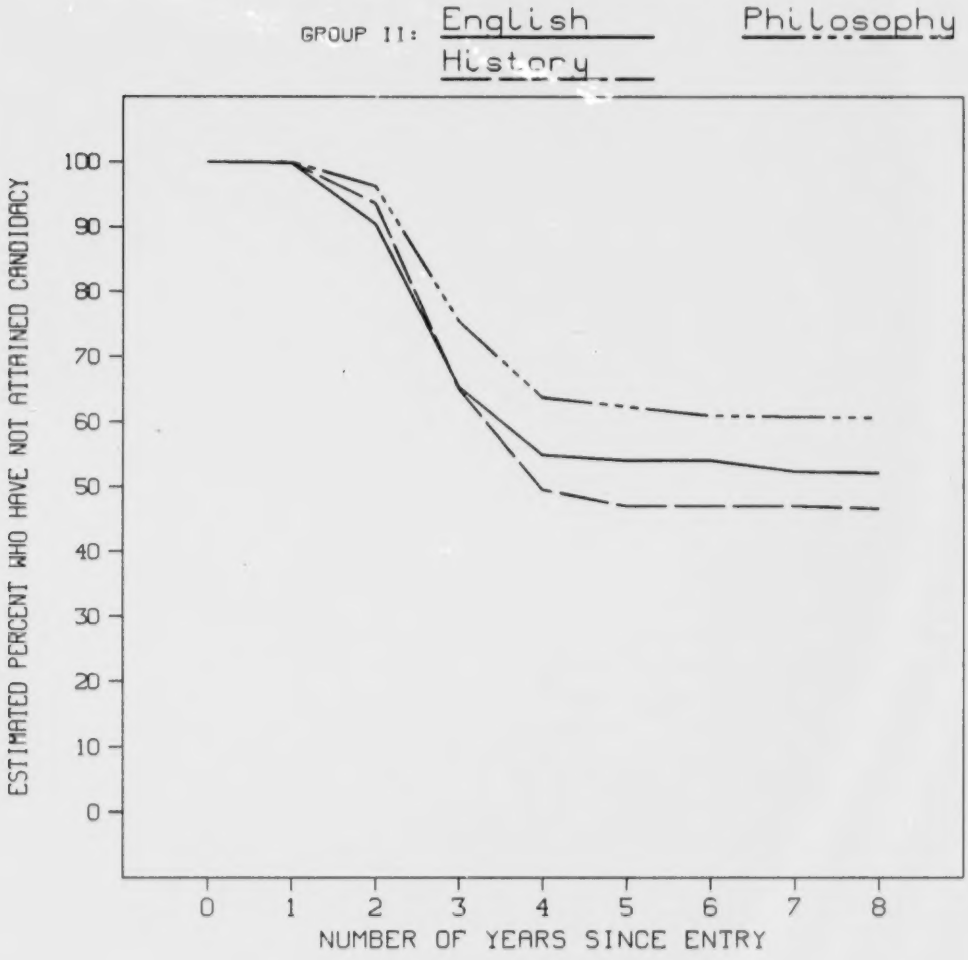
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 3



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 8

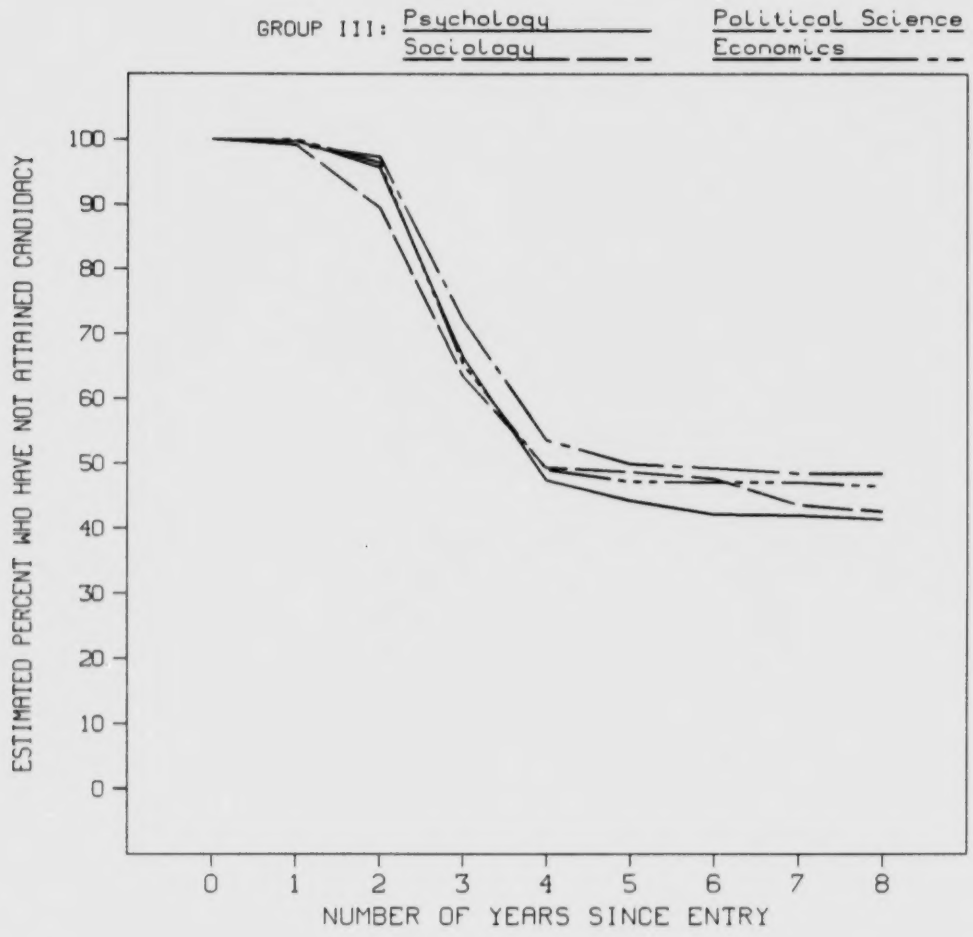
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 3



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 9

BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 3



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

Table 4

Estimated Percentages of Students Achieving Ph.D. Candidacy by Department[a]

	Five Years After Entry			Eight Years After Entry		
	School 1	School 2	School 3	School 1	School 2	School 3
Group I						
Chemistry	70	63	73	71	63	73
Physics	56	45	53	81	49	54
Mathematics	61	49	69	64	51	70
Computer Science	56	36	44	69	44	48
Group II						
English	47	42	46	60	49	48
Philosophy	45	29	38	50	30	39
History	54	50	53	62	50	53
Group III						
Psychology	61	62	56	78	62	59
Political Science	56	45	53	64	47	54
Sociology	43	38	51	68	49	57
Economics	46	34	50	67	41	52

[a]The tabled quantities are $100(1 - \hat{S}(x))$, where x is the number of years since entry. Estimates of candidacy and graduation rates are based on slightly different groups of students. See section on survival analysis.

Graduation Analysis. Figures 10-18 show the survival functions for the graduation analyses for each of the three groups of departments at School 1, School 2, and School 3. Table 5 shows the estimated percentages of students completing the Ph.D. five and eight years after entry for the 11 departments in the three schools.[6] The results show that School 3 had the highest graduation rates in every department five years after entry, whereas School 1 nearly always had the highest rates at eight years. In about half the departments at School 1 and most departments at the two other schools, the estimated percentage of students graduating by year 8 was less than 50.

At School 3, graduation was most likely in years 5 and 6 in all departments. By year 7, the survival functions had leveled out. At School 1 and School 2, by contrast, the survival functions were still decreasing at year 8. In fact, in these two schools, the hazard functions for all Group II departments were increasing at year 8, suggesting that the most likely year of graduation for these departments was beyond year 8. The hazard functions for Political Science and Sociology at School 1 and for Computer Science, Political Science, and Economics at School 2 were also increasing at year 8; the hazards peaked during years 5 through 7 for the remaining departments at these two schools.

In general, the results in Table 5 revealed more similarities within disciplines than within schools. Group I and Group III departments were typically more diverse than Group II. Within Group I, Chemistry always showed the highest graduation rates. Graduation rates were very similar for the three departments in Group II. In Group III, Psychology always had the highest graduation rates, followed by Economics. (Economics and Political Science had identical rates at year 5 at School 2.) In general, Group I departments had the highest graduation rates, followed by Group III and then Group II, paralleling the ordering observed for the candidacy analysis.

Analyses for Ethnic and Gender Groups in Department Clusters

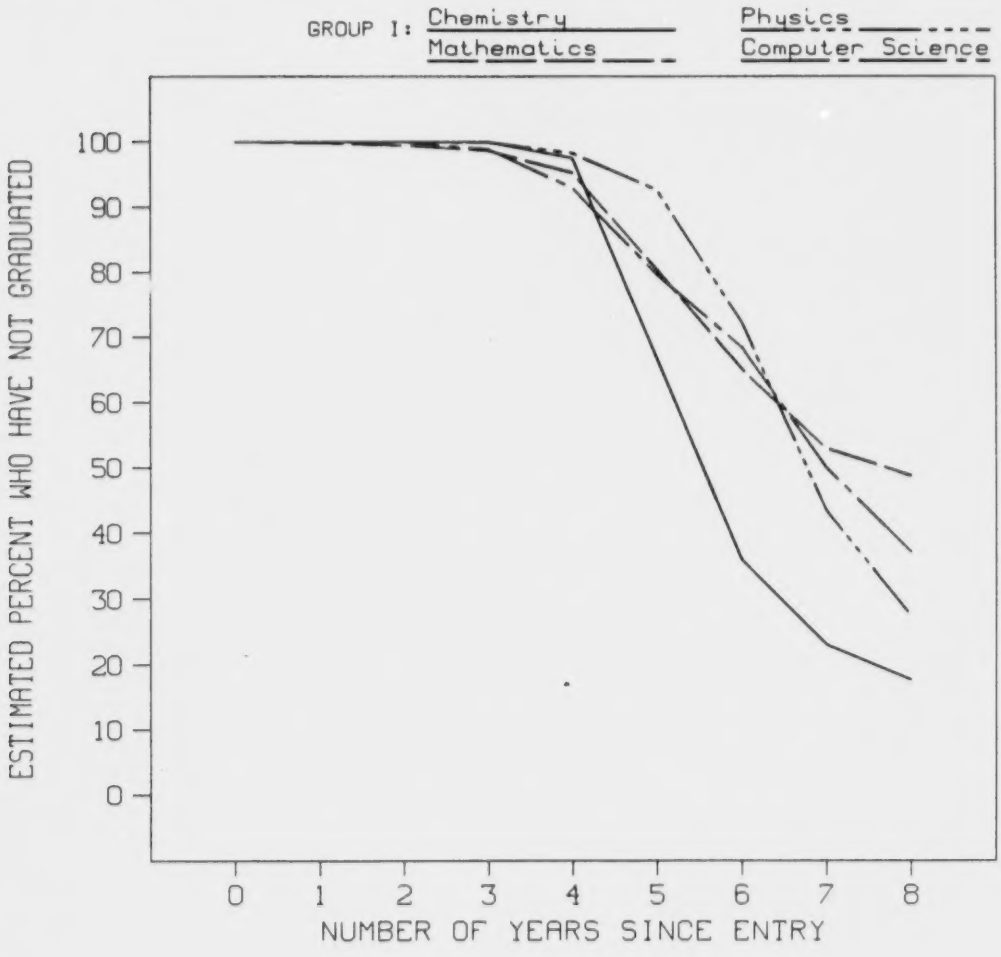
Because of the small number of minority students, comparisons of attainment patterns for ethnic groups within individual departments could not be conducted. Instead, for purposes of conducting survival analyses for ethnic and gender groups, departments that had similar survival curves in the 11-department analyses were combined to form clusters. The results of the 11-department analyses of candidacy and graduation were examined, focusing on years 5 and 8 (see Tables 4 and 5) in order to find groups of departments with survival curves that were close together in both analyses in all three schools. Two departmental clusters were derived that met this criterion reasonably well: Cluster 1, which includes Mathematics, Physics, Computer Science, Economics, and Psychology, and Cluster 2, which includes

---

[6]This study focuses on the elapsed time between entering graduate school and attaining the Ph.D. Other studies focus on the elapsed time between the baccalaureate and the doctorate or on the amount of registered time in graduate school.

FIGURE 10

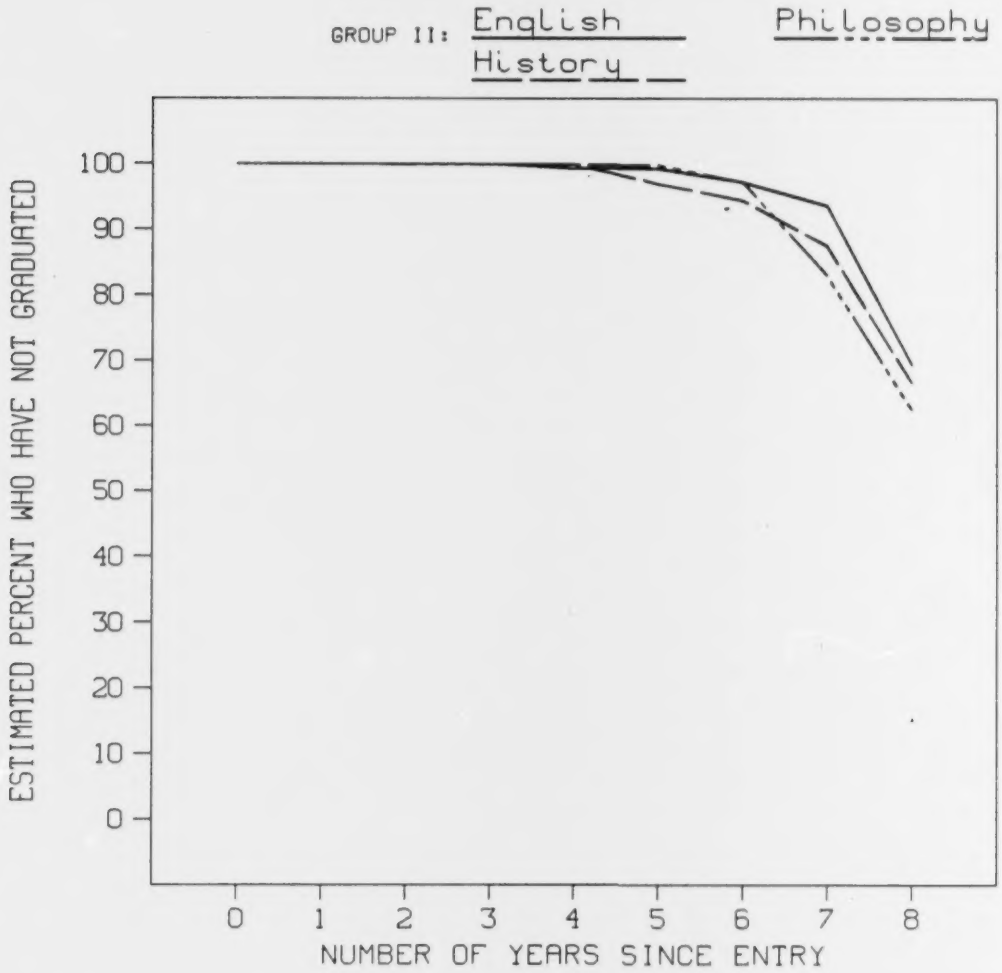
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 11

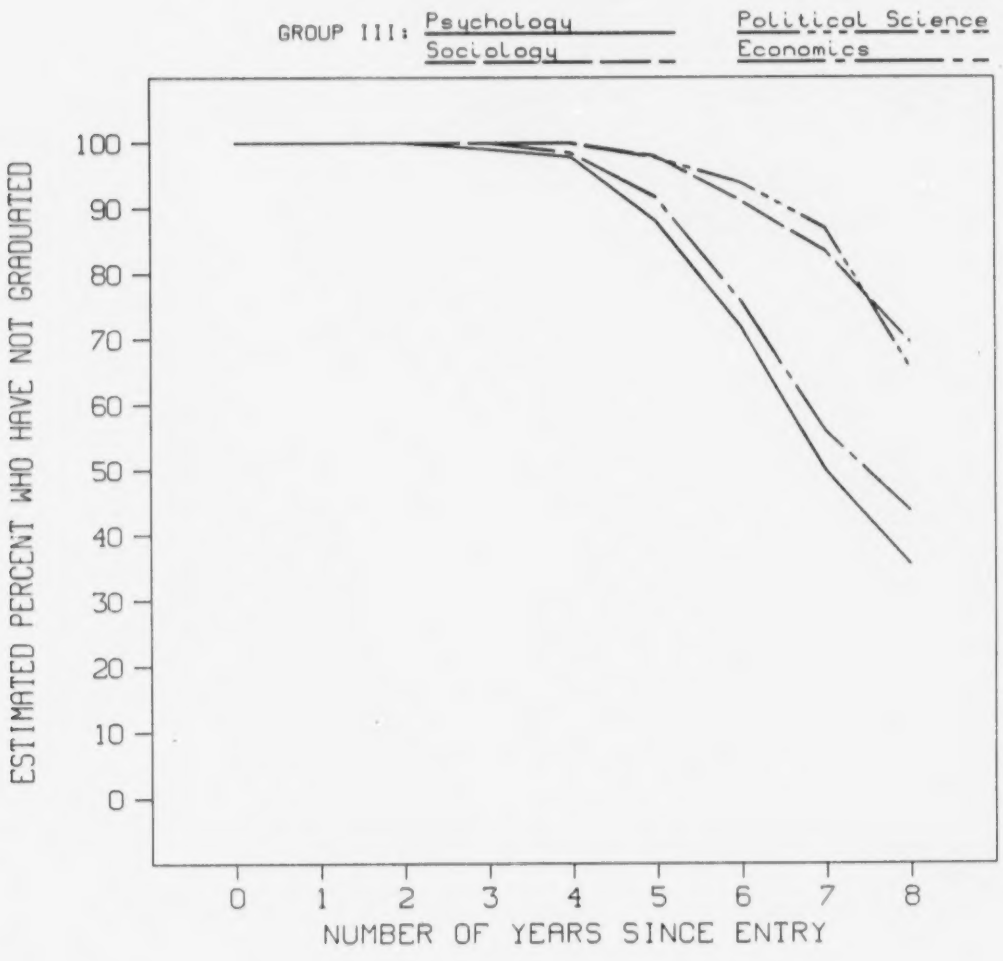
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 12

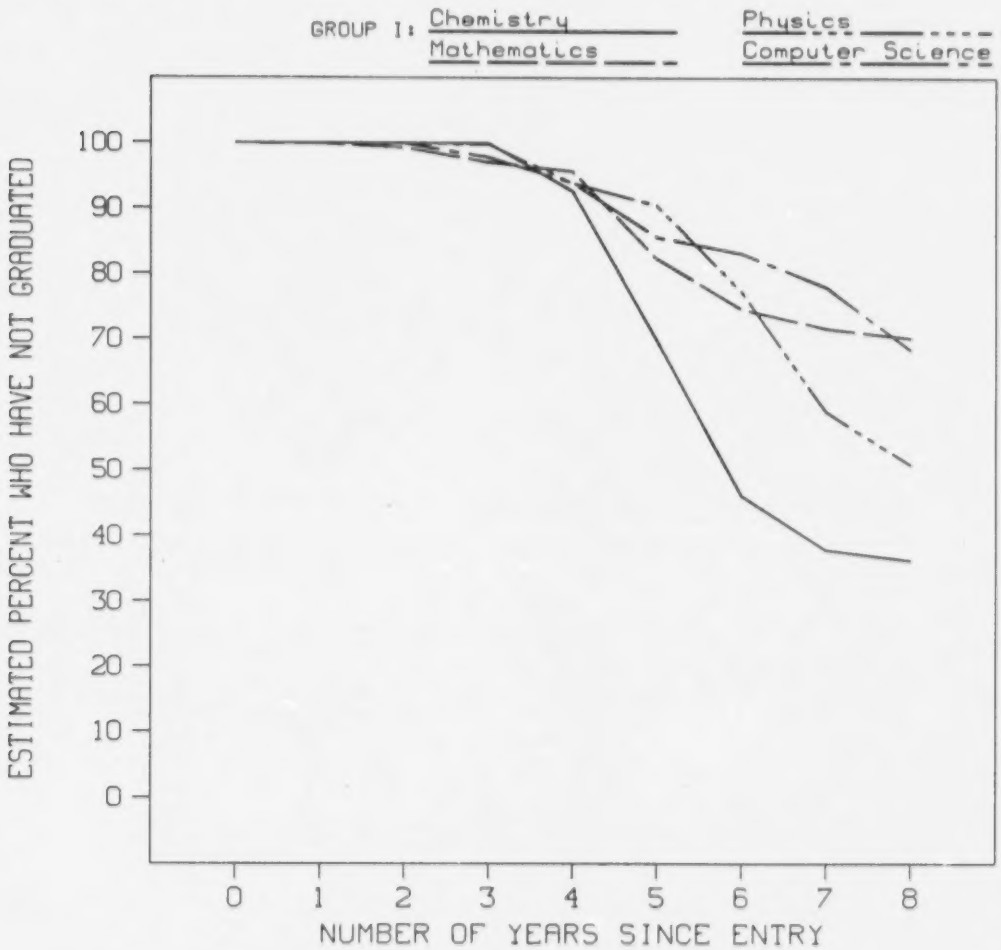
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 13

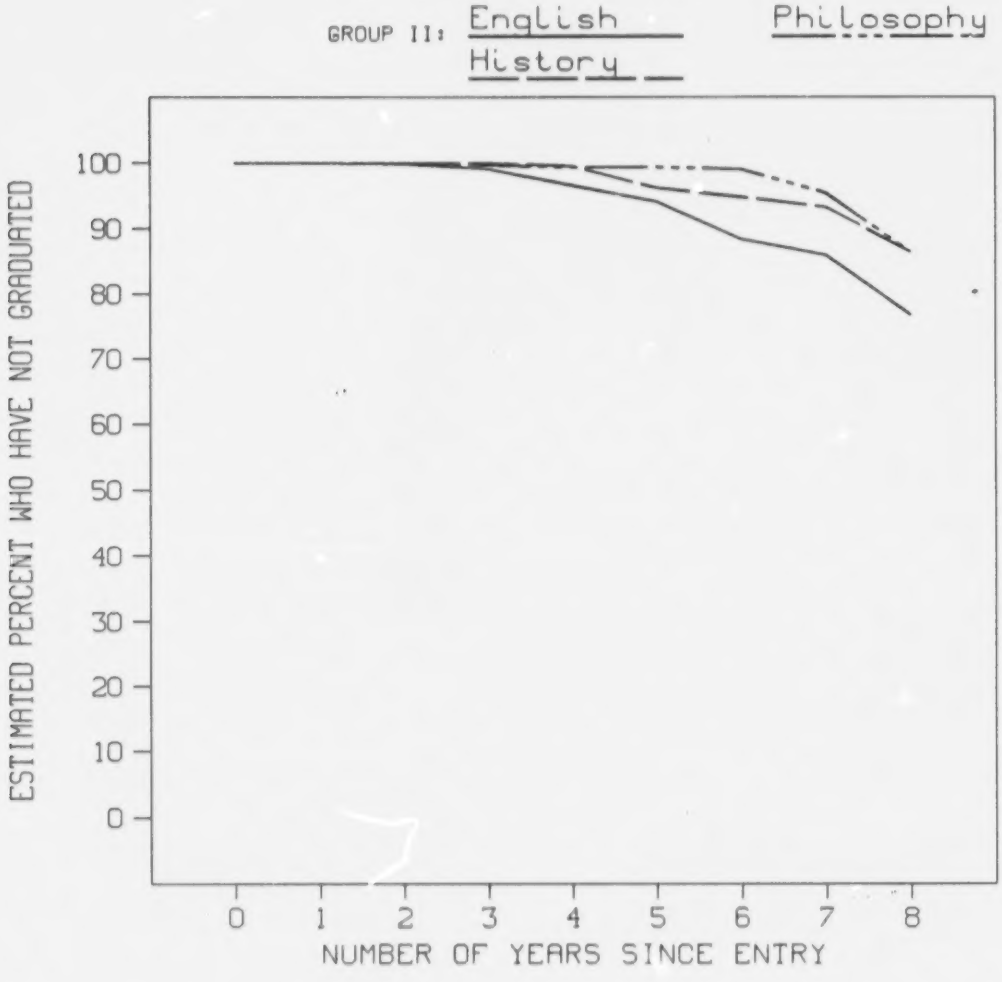
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 14

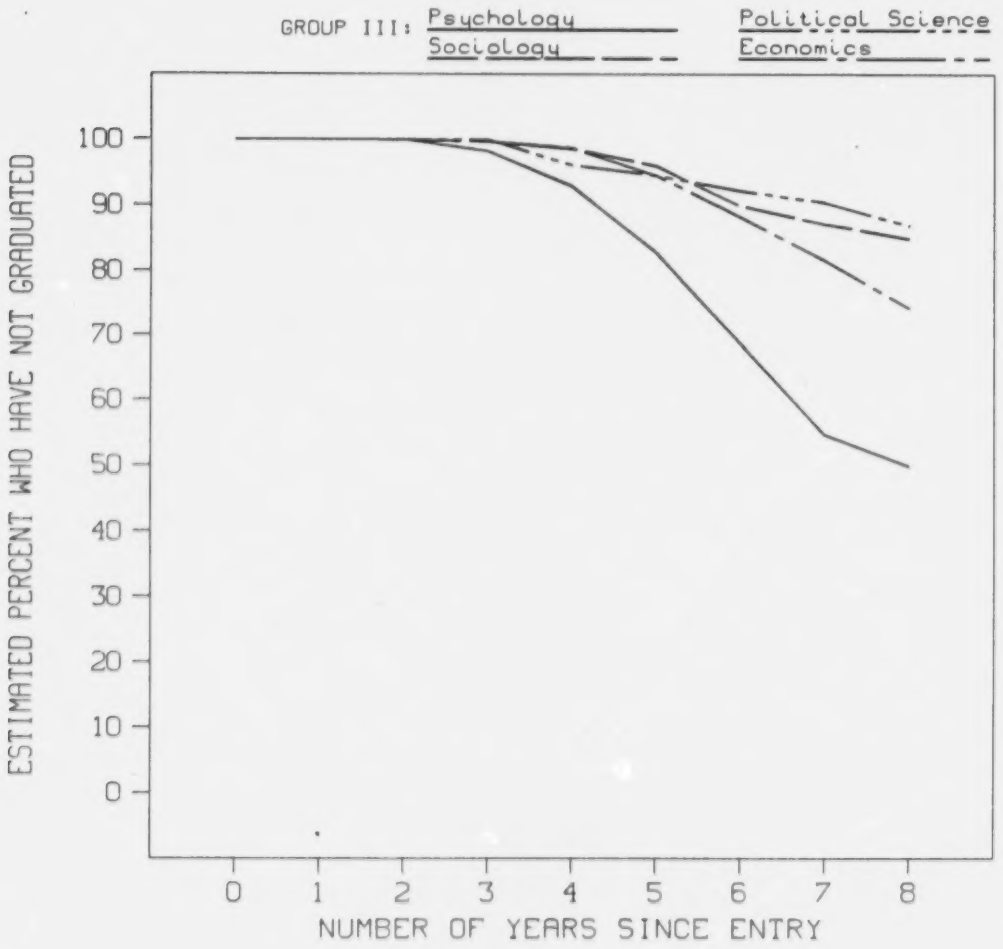
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 15

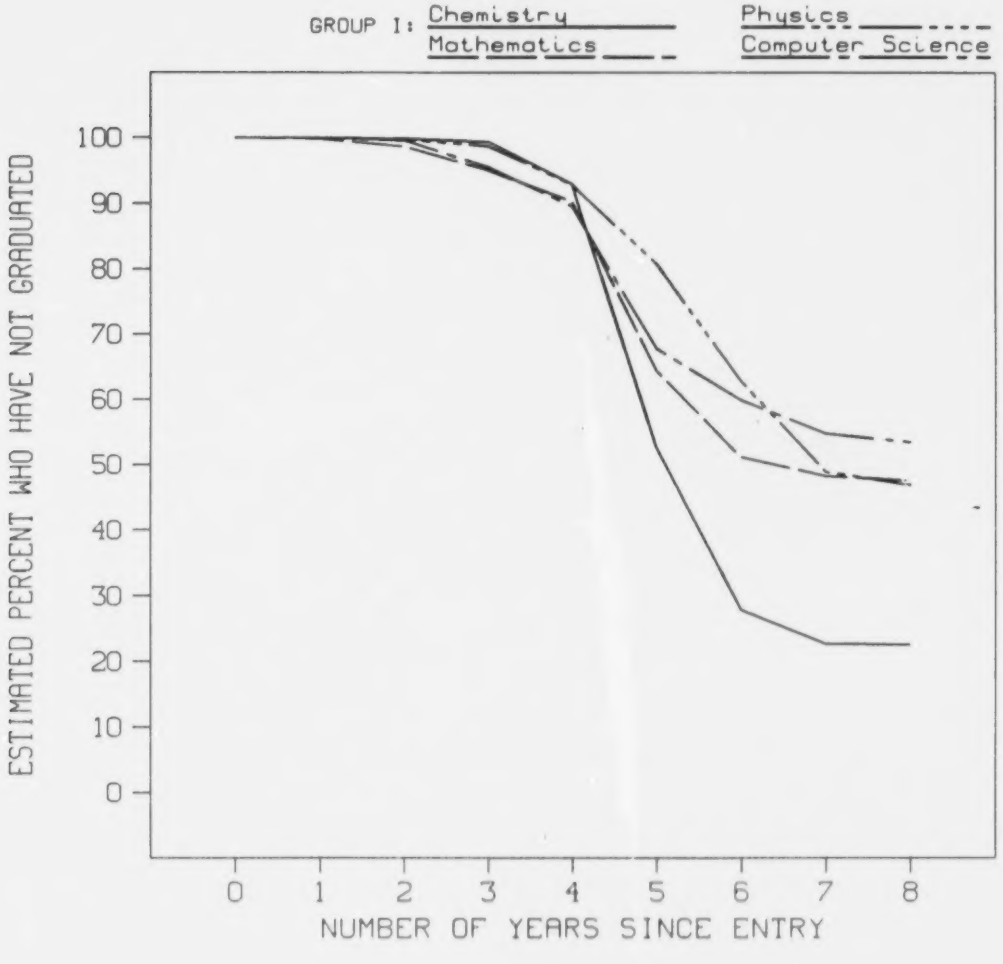
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 16

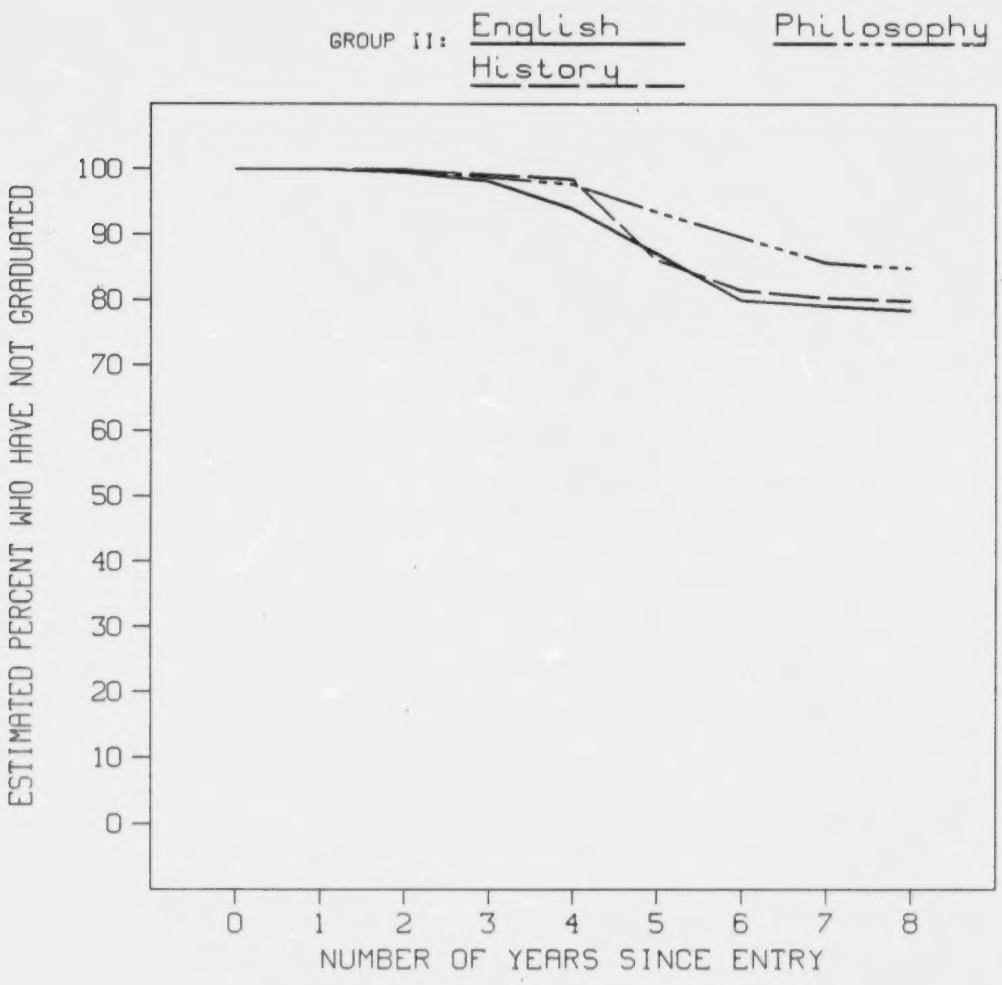
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 3



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 17

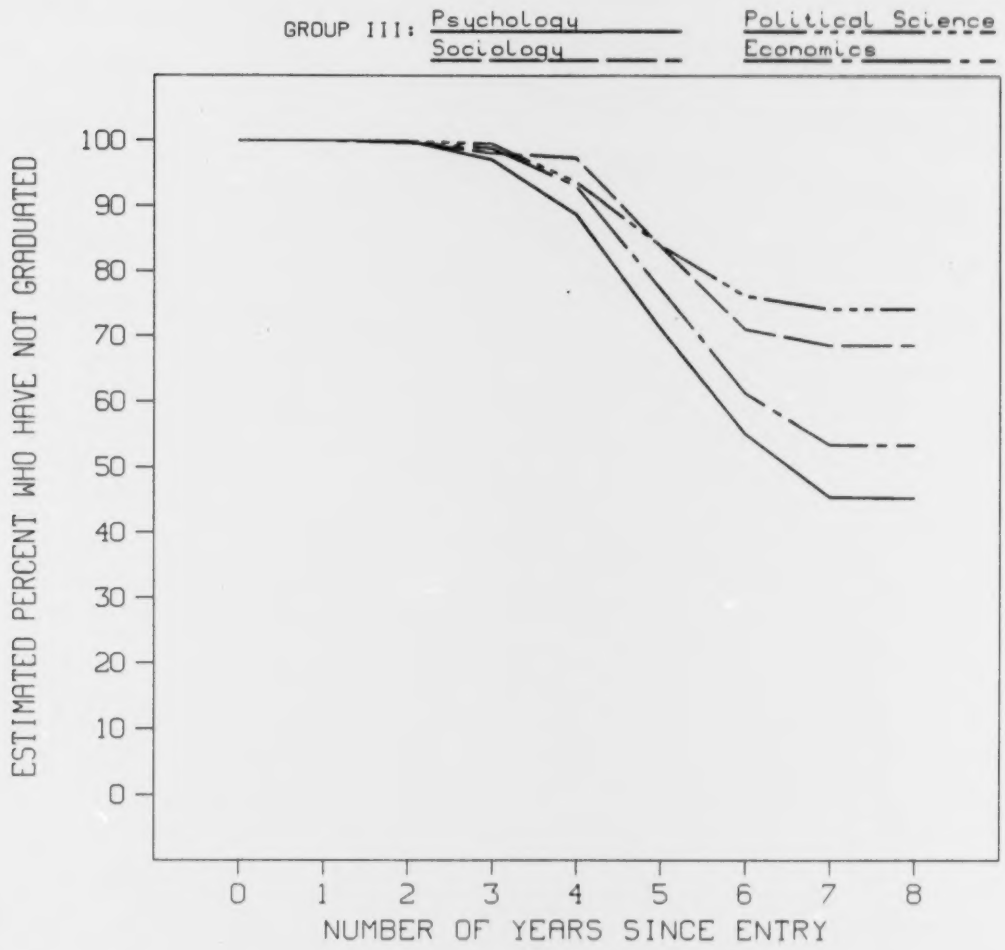
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 3



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 18

BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 3



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

Table 5

Estimated Percentages of Students Completing Ph.D. Degrees by Department[a]

	Five Years After Entry			Eight Years After Entry		
	School 1	School 2	School 3	School 1	School 2	School 3
Group I						
Chemistry	33	30	47	82	64	77
Physics	7	10	19	73	49	53
Mathematics	20	18	36	51	30	52
Computer Science	20	14	32	63	32	46
Group II						
English	1	6	13	31	23	22
Philosophy	0	1	7	38	14	15
History	3	4	14	33	14	20
Group III						
Psychology	12	17	29	65	50	55
Political Science	2	5	16	34	13	26
Sociology	2	4	16	31	15	31
Economics	8	5	23	56	26	47

[a]The tabled quantities are $100(1 - \hat{S}(x))$, where x is the number of years since entry. Estimates of candidacy and graduation rates are based on slightly different groups of students. See section on survival analysis.

Philosophy, History, English, Sociology, and Political Science. Chemistry did not fit well in either cluster and was therefore excluded from the ethnic and gender group analyses. Except for Psychology, it is clear that the departments in Cluster 1 are quantitatively oriented departments. Although it may seem odd that Psychology fell into this group, a similar finding was obtained in a survival analysis conducted at Stanford University (Mathematical Methods in Education, 1983), in which Psychology was found to group naturally with Chemistry and Physics. The second departmental cluster includes both social science and humanities departments.

At each school, survival analysis results were obtained within each cluster for men and women and for ethnic groups with sufficiently large samples. At School 1, results were obtained for Asian, Black, Hispanic, White, and foreign students; at School 2, results were obtained for White and foreign students only; and at School 3, results were obtained for Black, White, and foreign students. Results for Asian, Black, and Hispanic students in general, and for foreign students in Cluster 2 at School 2 are not as well determined as those for the larger groups in the analysis. (In particular, the number of Black students in Cluster 1 at School 3 was only seven.) Because of sample size limitations, results were not obtained for men and women within ethnic categories.

Ethnic Groups--Candidacy. Results of the candidacy analyses for ethnic groups are given in Figures 19-24. Within each school, results are graphed separately for Clusters 1 and 2. Estimated percentages of students receiving candidacy five and eight years after entry are given in Table 6 for each of the included ethnic groups in Clusters 1 and 2 at the three schools. Sample sizes are given in Table D-14.

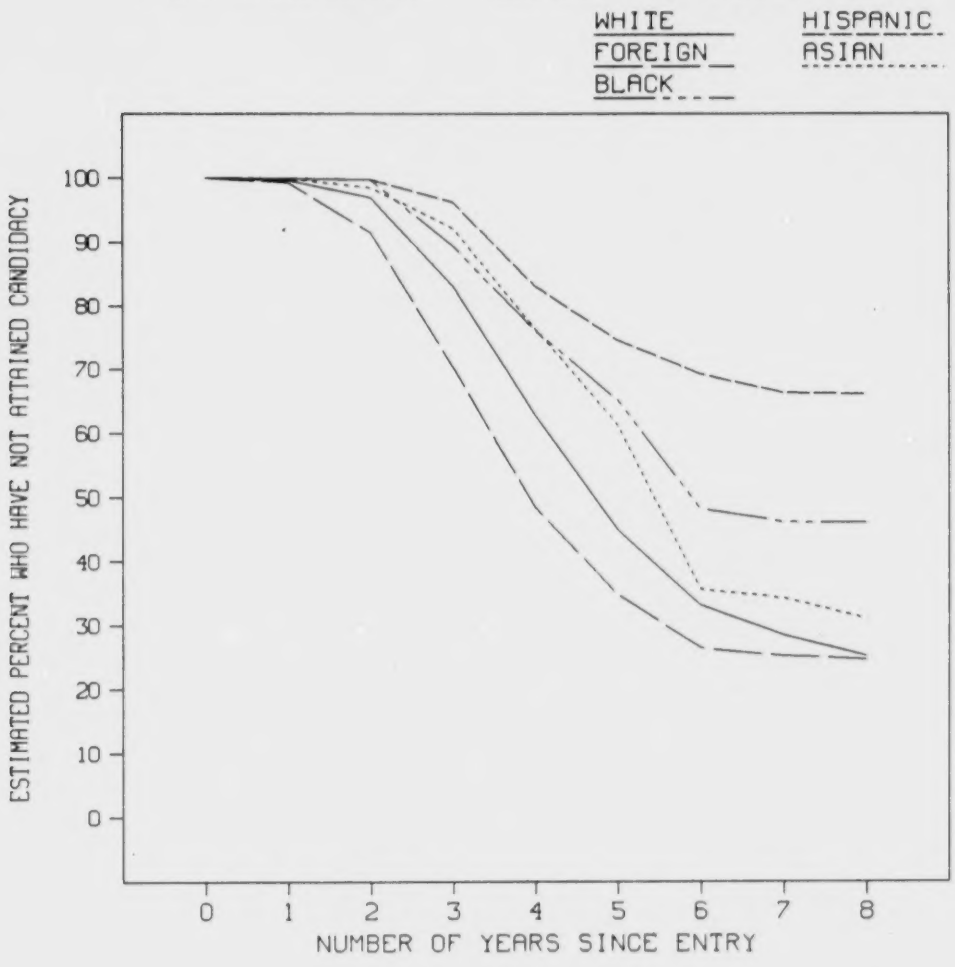
The four plots for School 1 and School 3 show a strikingly similar pattern: In each case, foreign students had consistently higher candidacy rates than did White American students, who, in turn had higher rates than Black Americans. (As noted above, in Cluster 1 at School 3, the results for Blacks are based on only seven students.) At School 1, Asian Americans had candidacy rates between those of Whites and Blacks; Hispanics had the lowest rates. The results at School 2 were somewhat different: In Cluster 1, the quantitative group of departments, White and foreign students had nearly identical rates, whereas in Cluster 2, White Americans had higher candidacy rates than foreign students.

Ethnic Groups--Graduation. Results of the graduation analyses for ethnic groups are given in Figures 25-30. Within each school, results are graphed separately for Clusters 1 and 2. Estimated percentages of students completing Ph.D. degrees five and eight years after entry are given in Table 7 for each of the included ethnic groups in Clusters 1 and 2 at the three schools. Sample sizes are given in Table D-15.

The plots for School 3 show the same ordering as the candidacy plots: foreign students had the highest graduation rates, followed by White Americans; Black Americans had attainment rates equal to or close to zero.

FIGURE 19

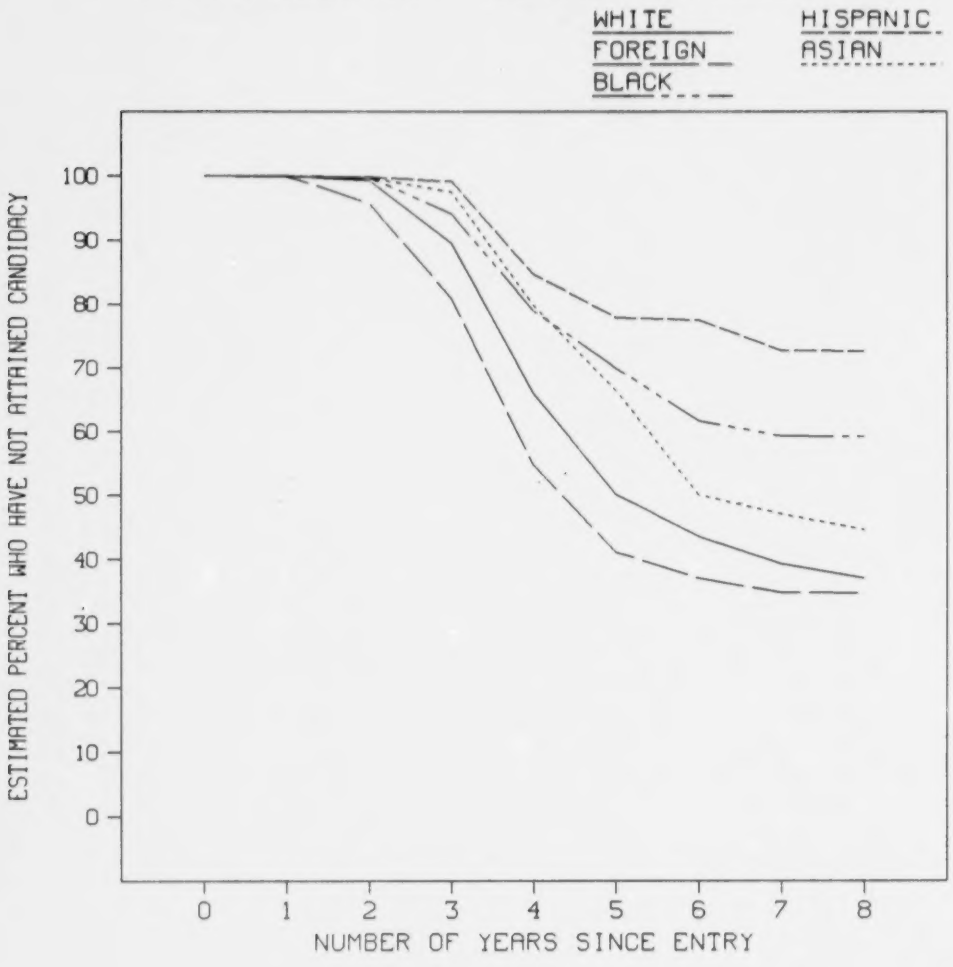
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 1, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 20

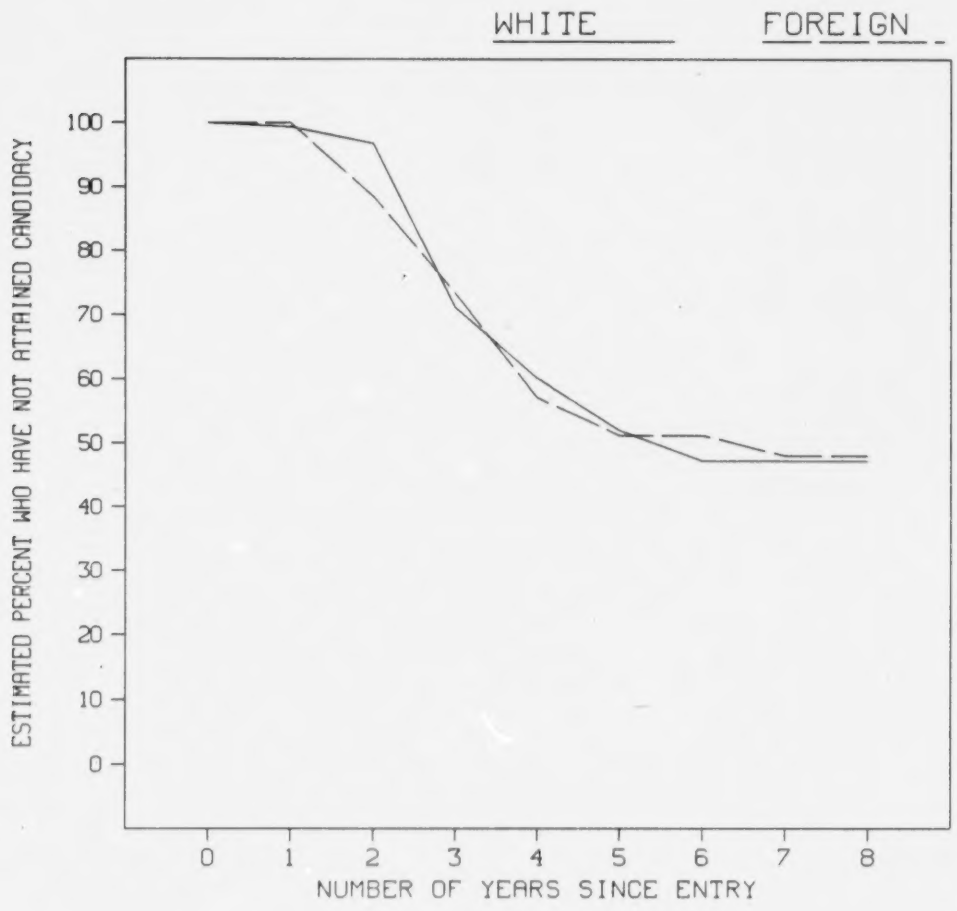
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 1, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 21

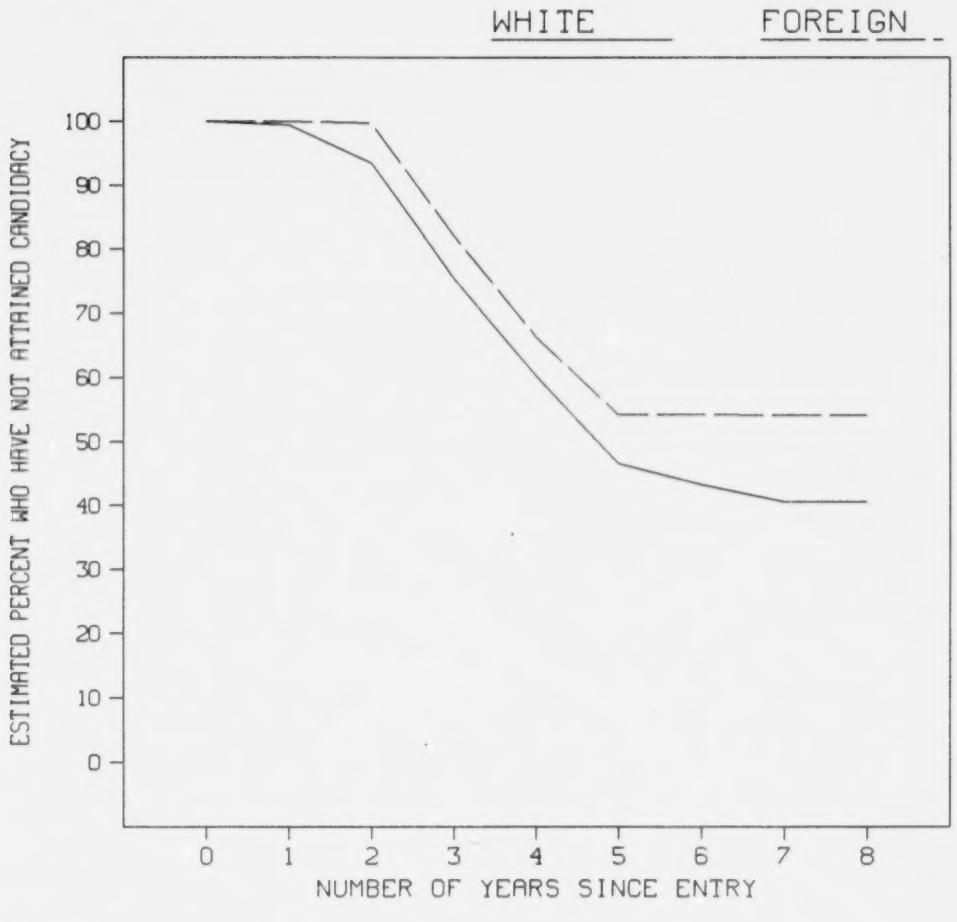
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 2, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 22

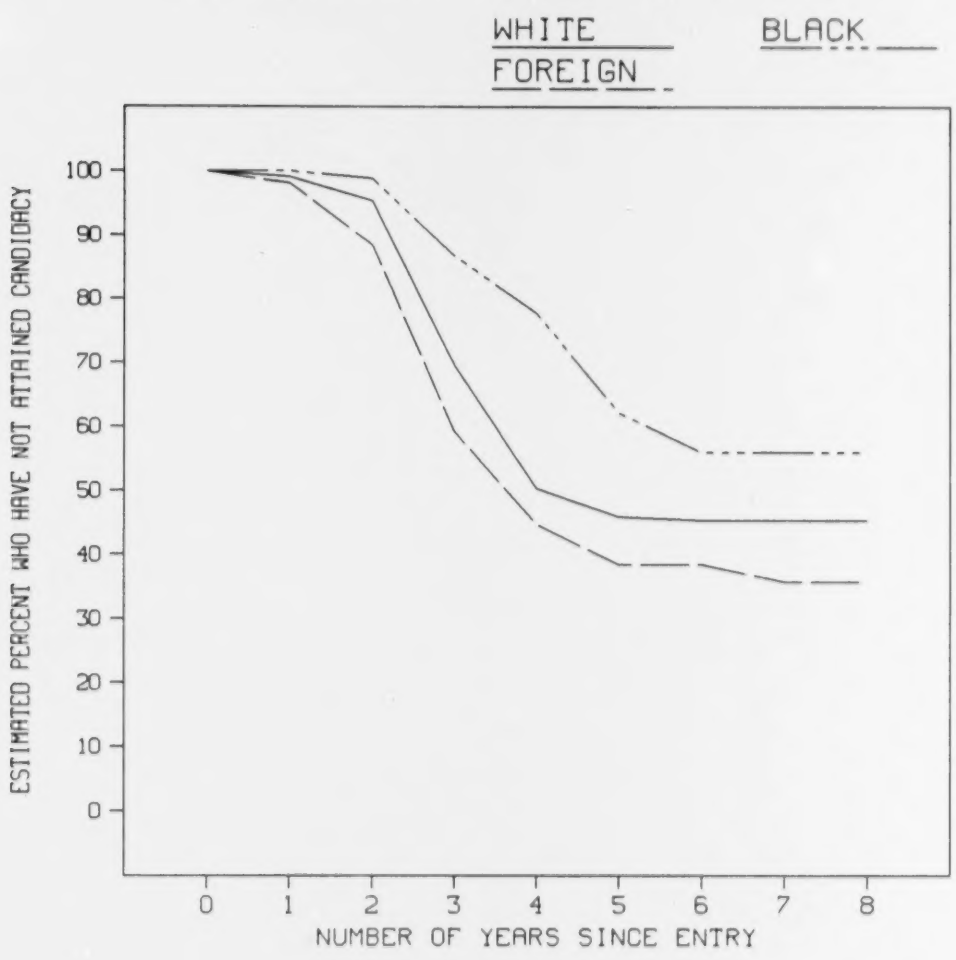
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 2, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 23

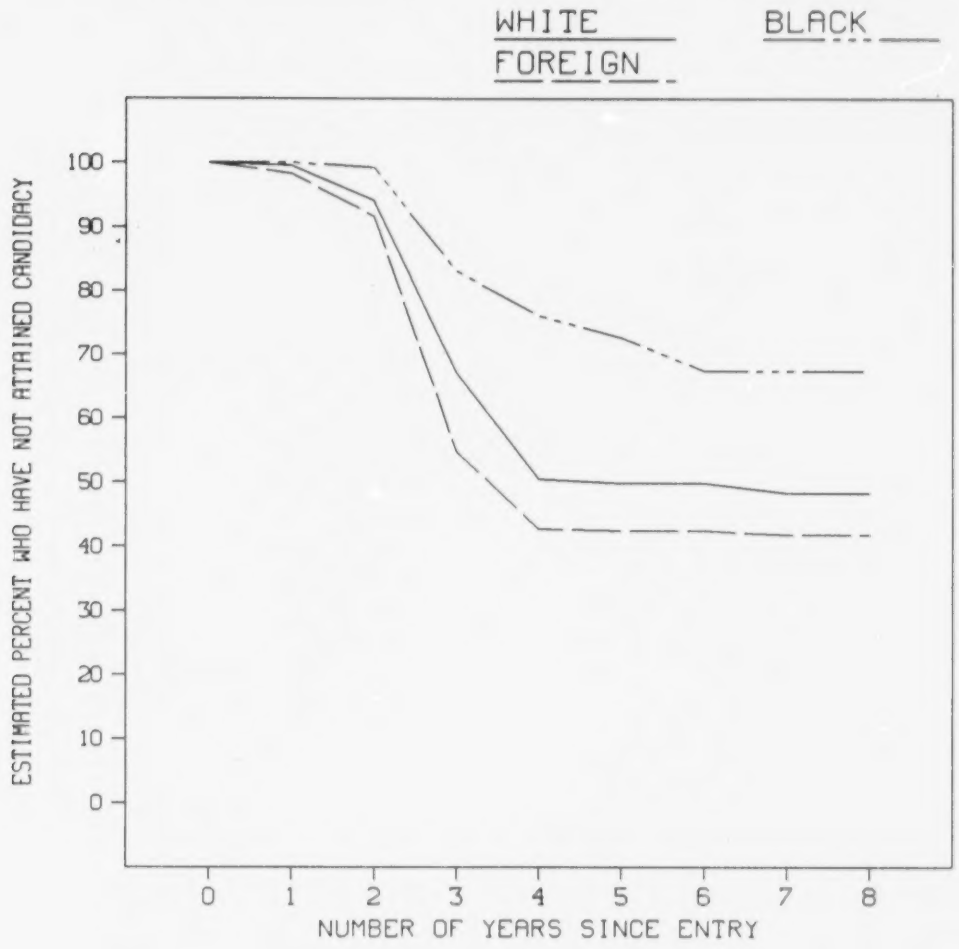
BAYES ESTIMATES OF SURVIVAL FUNCTIONS FOR CANDIDACY: SCHOOL 3, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 24

BAYES ESTIMATES OF SURVIVAL FUNCTIONS FOR CANDIDACY: SCHOOL 3, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

Table 6

Estimated Percentages of Students Achieving Ph.D. Candidacy by Ethnic Group[a]

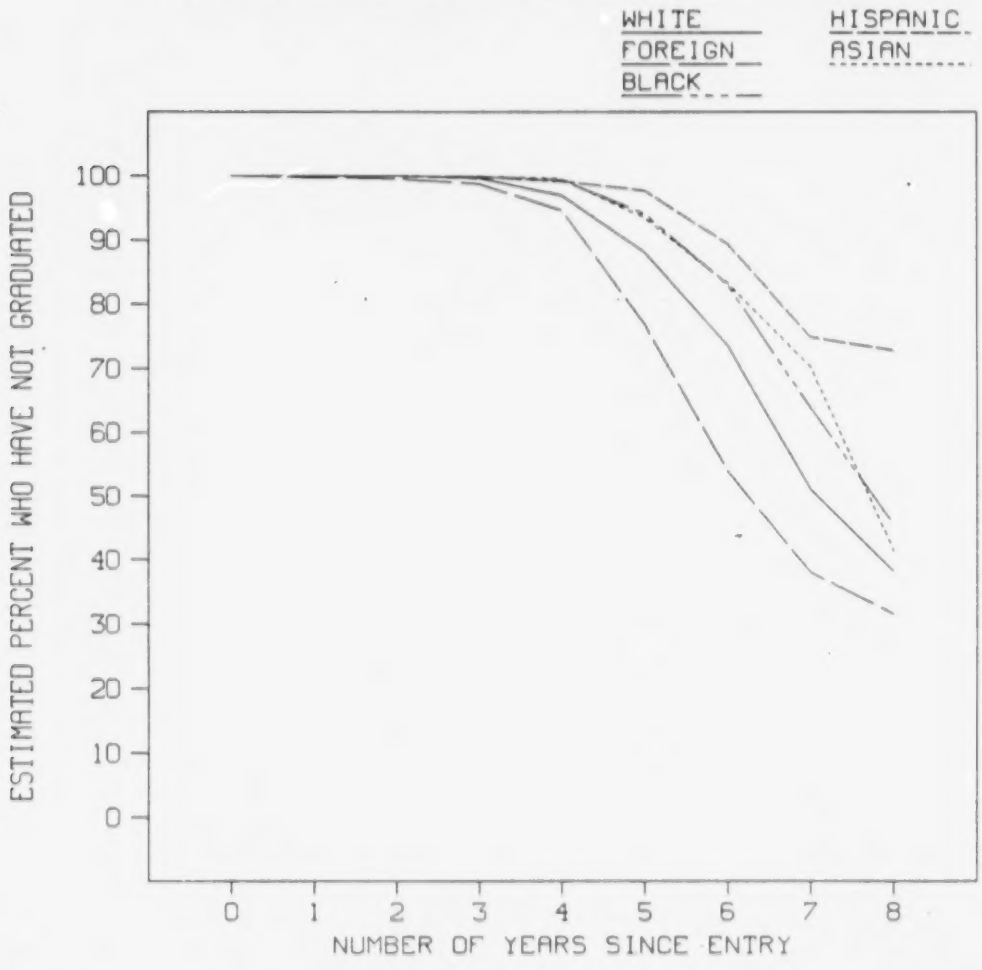
	Five Years After Entry			Eight Years After Entry		
	School 1	School 2	School 3	School 1	School 2	School 3
Cluster 1[b]						
Asian	39			69		
Black	35		38	54		44
Hispanic	26			34		
White	55	48	54	75	53	55
Foreign	65	49	62	75	52	64
Cluster 2						
Asian	34			55		
Black	30		27	41		33
Hispanic	22			27		
White	50	53	50	63	59	52
Foreign	59	46	58	65	46	58

[a]The tabled quantities are $100(1 - \hat{S}(x))$, where x is the number of years since entry. Estimates of candidacy and graduation rates are based on slightly different groups of students. See section on survival analysis.

[b]Cluster 1 includes Mathematics, Physics, Computer Science, Economics, and Psychology. Cluster 2 includes Philosophy, History, English, Sociology, and Political Science.

FIGURE 25

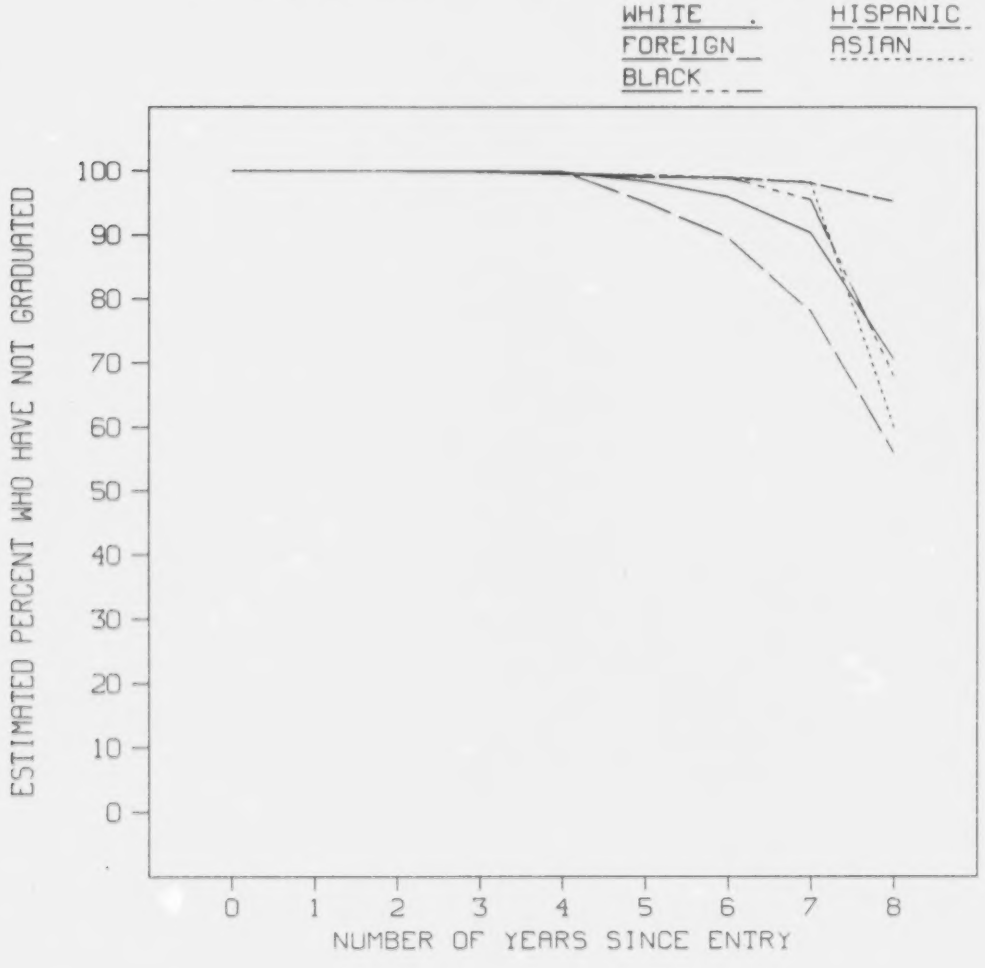
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 1, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 26

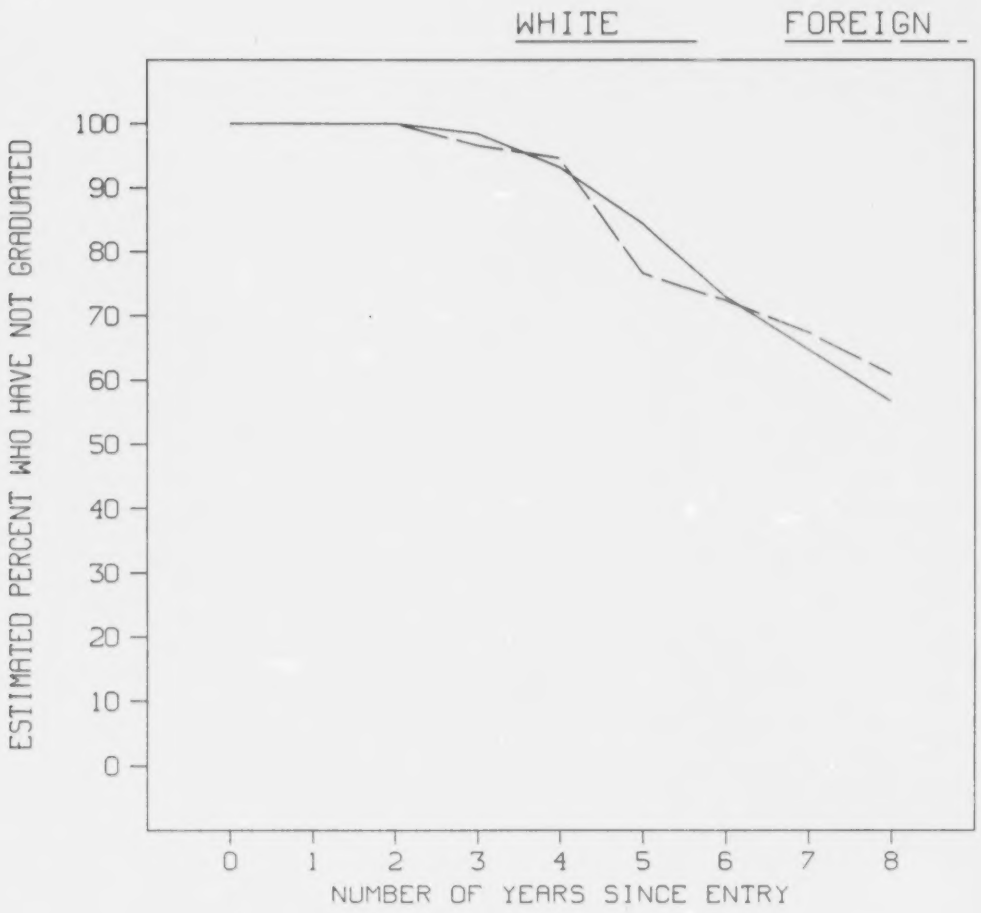
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 1, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 27

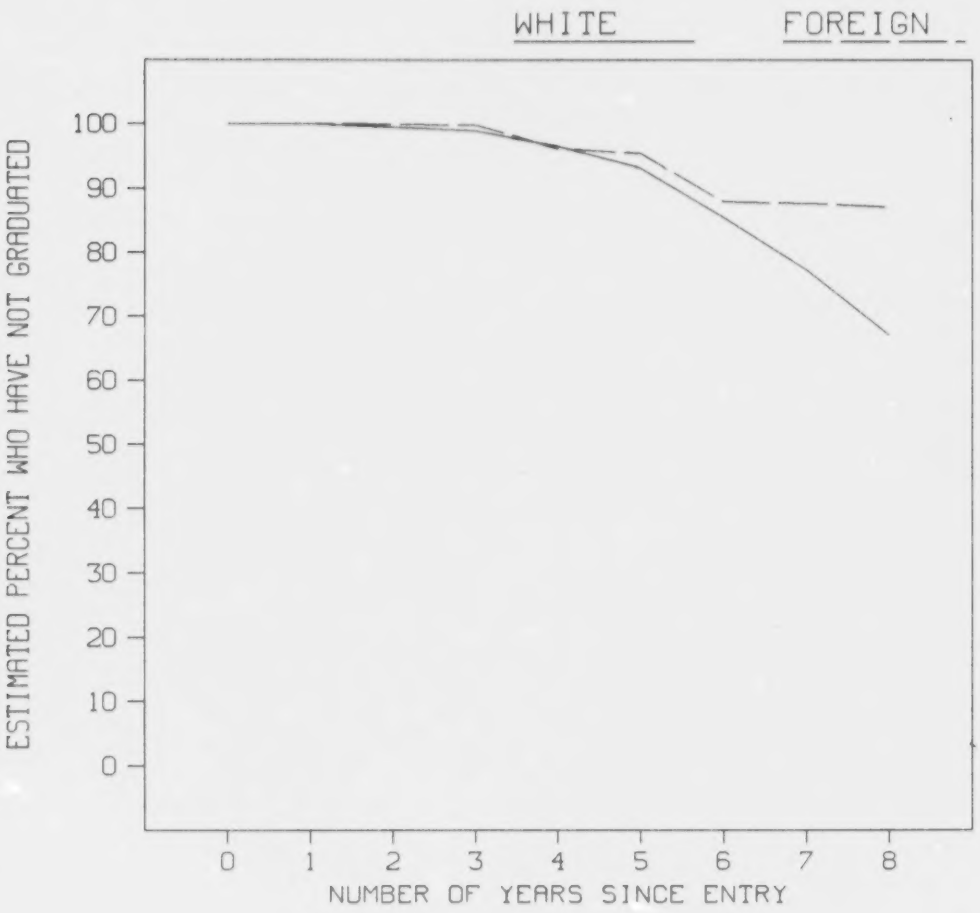
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 2, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 28

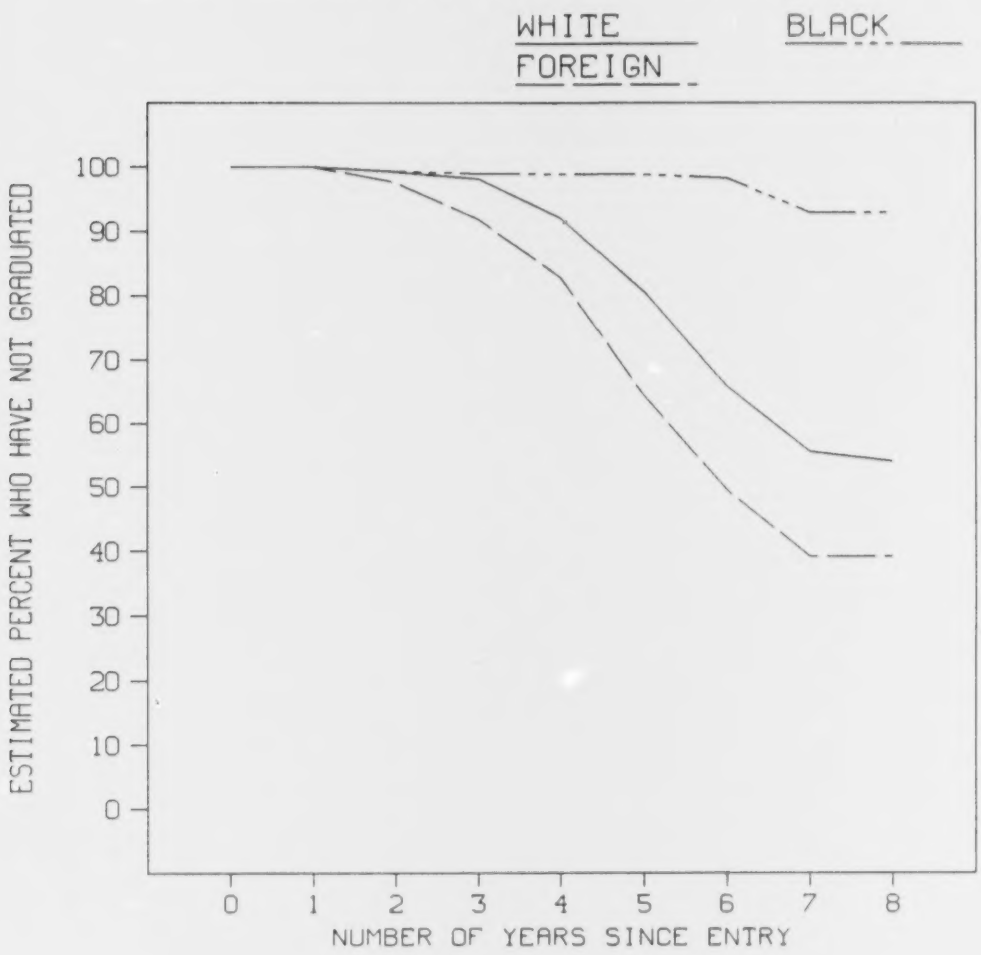
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 2, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 29

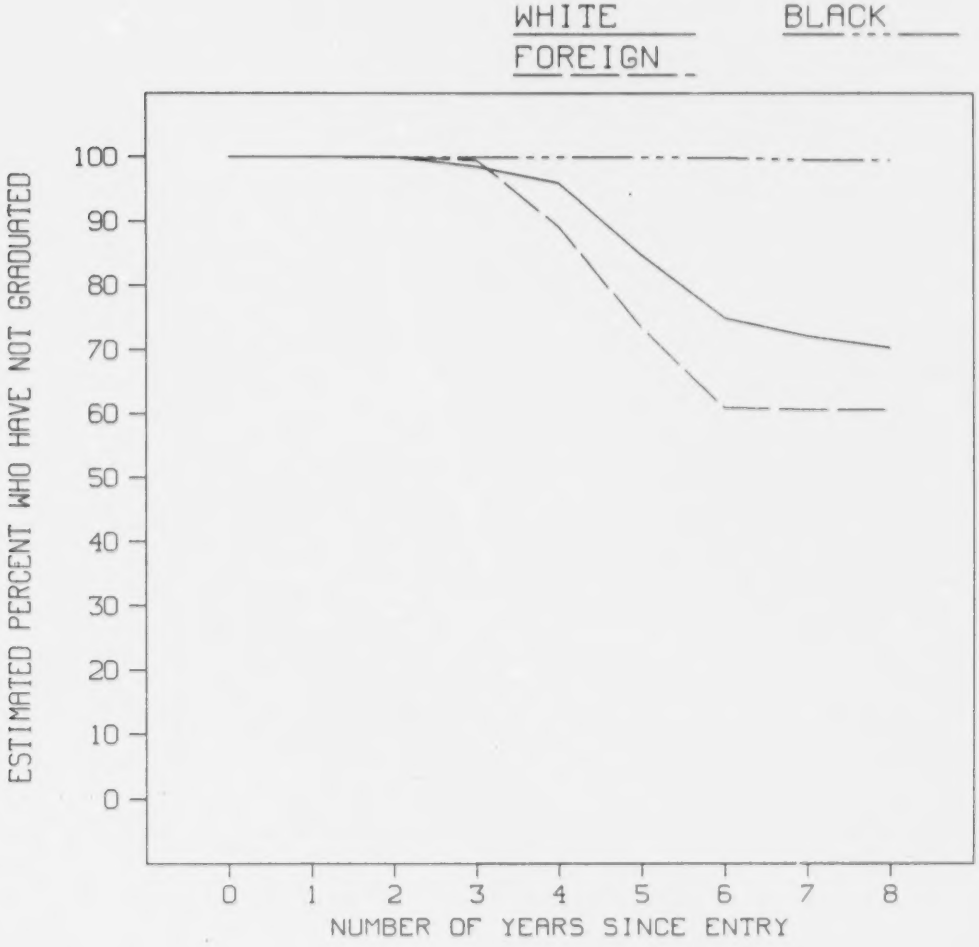
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 3, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 30

BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 3, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

Table 7

Estimated Percentages of Students Completing Ph.D. Degrees by Ethnic Group[a]

	Five Years After Entry			Eight Years After Entry		
	School 1	School 2	School 3	School 1	School 2	School 3
Cluster 1[b]						
Asian	7			59		
Black	6		1	54		7
Hispanic	2			27		
White	12	16	19	62	43	46
Foreign	23	23	36	68	39	61
Cluster 2						
Asian	1			40		
Black	1		0	32		1
Hispanic	1			5		
White	2	7	15	29	33	30
Foreign	5	5	27	44	13	39

[a]The tabled quantities are $100(1 - \hat{S}(x))$, where x is the number of years since entry. Estimates of candidacy and graduation rates are based on slightly different groups of students. See section on survival analysis.

[b]Cluster 1 includes Mathematics, Physics, Computer Science, Economics, and Psychology. Cluster 2 includes Philosophy, History, English, Sociology, and Political Science.

(Again, note that the results for Blacks in Cluster 1 at School 3 are based on only seven students.) At School 1, the Cluster 1 results were similar to those obtained in the candidacy analysis, except that rates for Asians and Blacks were virtually indistinguishable. In Cluster 2, all five groups were closer together, although graduation rates were clearly highest for foreign students and lowest for Hispanics. At School 2, Whites and foreign students in Cluster 1 had similar attainment rates, as in the candidacy analysis; in Cluster 2, White students had substantially higher graduation rates than foreign students after year 6.

Gender Groups--Candidacy. Results of the survival analyses for candidacy for men and women are given in Figures 31-36. Within each school, results are shown separately for Clusters 1 and 2. Estimated percentages of students receiving candidacy five and eight years after entry are given in Table 8 for men and women in Clusters 1 and 2 at the three schools. Sample sizes are given in Table D-16. In general, candidacy rates tended to be lower for women than for men, although, at School 3, the survival curves were very close, particularly in Cluster 1. The differences between men and women in the percentage of students estimated to have achieved candidacy ranged from 0 to 13% at year 5 and from 1 to 6% at year 8. The most pronounced difference between men and women occurred in Cluster 2 at School 2 between years 5 and 8.

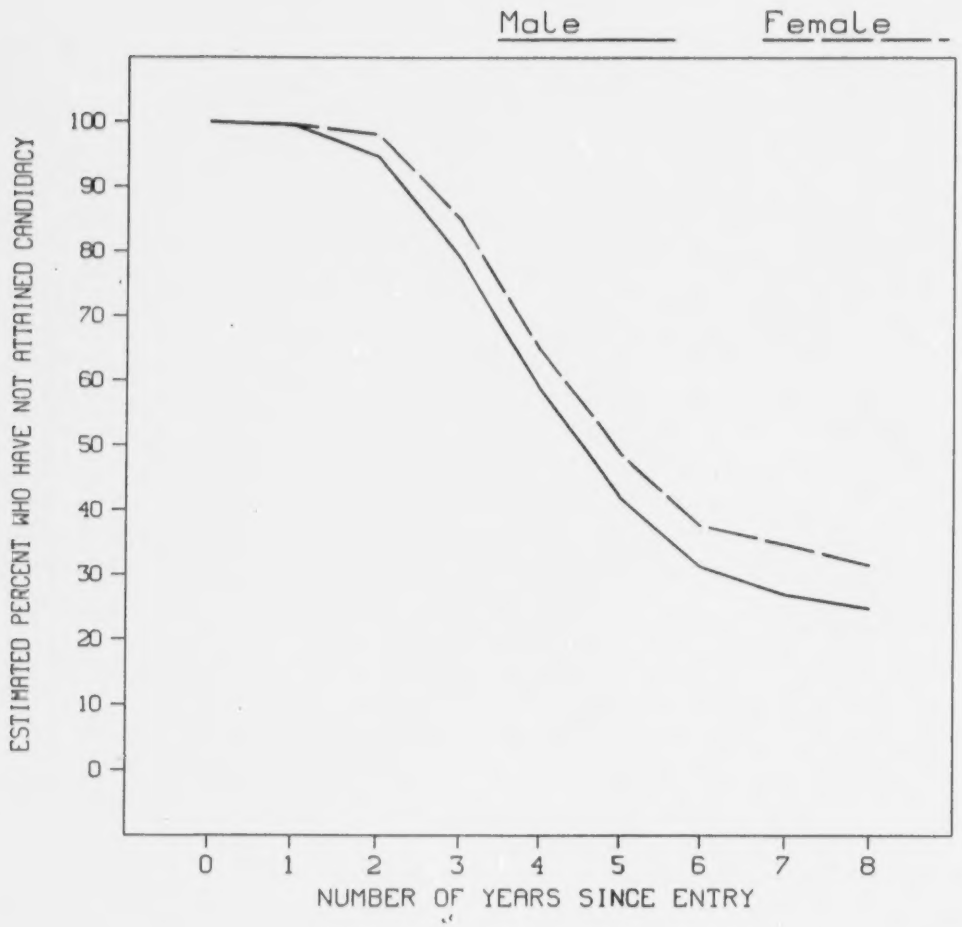
Gender Groups--Graduation. Results of the survival analysis for graduation for men and women are shown in Figures 37-42. Estimated percentages of students completing Ph.D. degrees five and eight years after entry are given in Table 9 for men and women in Clusters 1 and 2 at the three schools. Sample sizes are given in Table D-17. The results are similar to those obtained in the candidacy analysis in that attainment rates for men tended to be greater than those for women. Here, this pattern was evident for School 3 as well as the other two schools. Rates for men and women tended to be the most similar at School 1 and, again, the largest discrepancy occurred in Cluster 2 at School 2. The differences between men and women in the percentages of students estimated to have graduated ranged from 2 to 11% at year 5 and from 0 to 10% at year 8.

Relation of Candidacy and Graduation to Measures of Academic Potential

A correlational analysis, described below, was conducted to explore the association of candidacy and graduation with undergraduate grade-point average (UGPA), GRE verbal score (GREV), and GRE quantitative score (GREQ). As shown

FIGURE 31

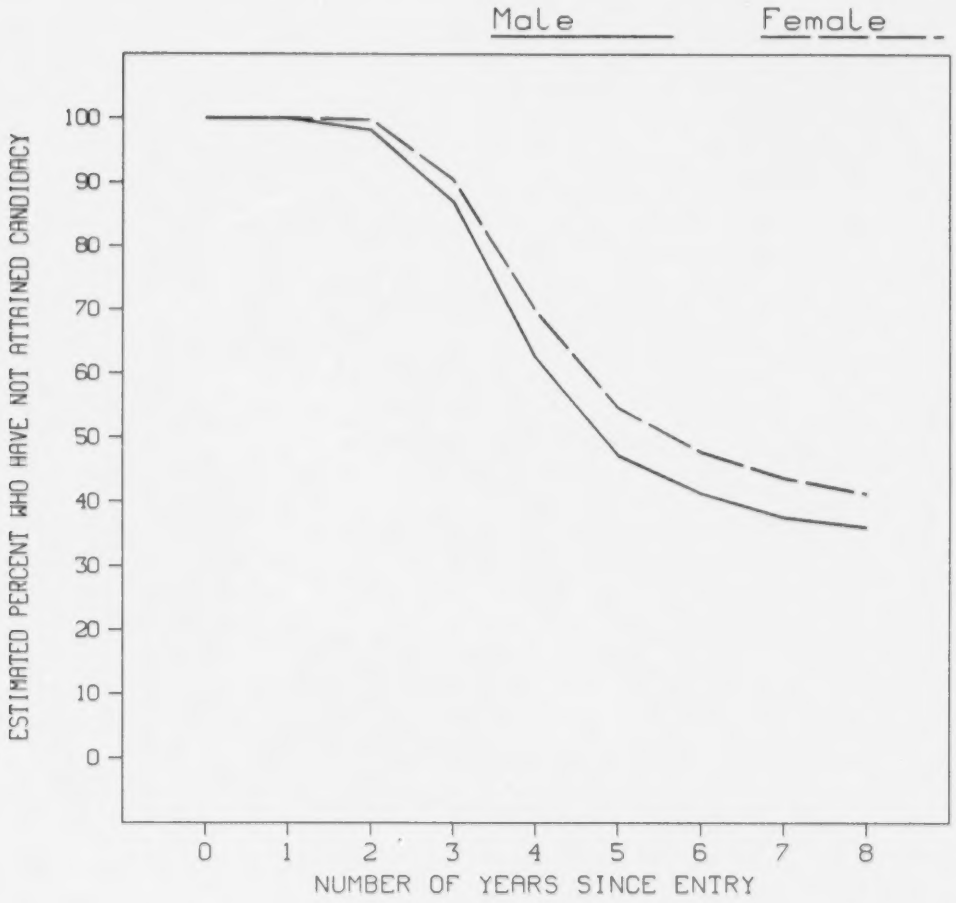
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 1, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 32

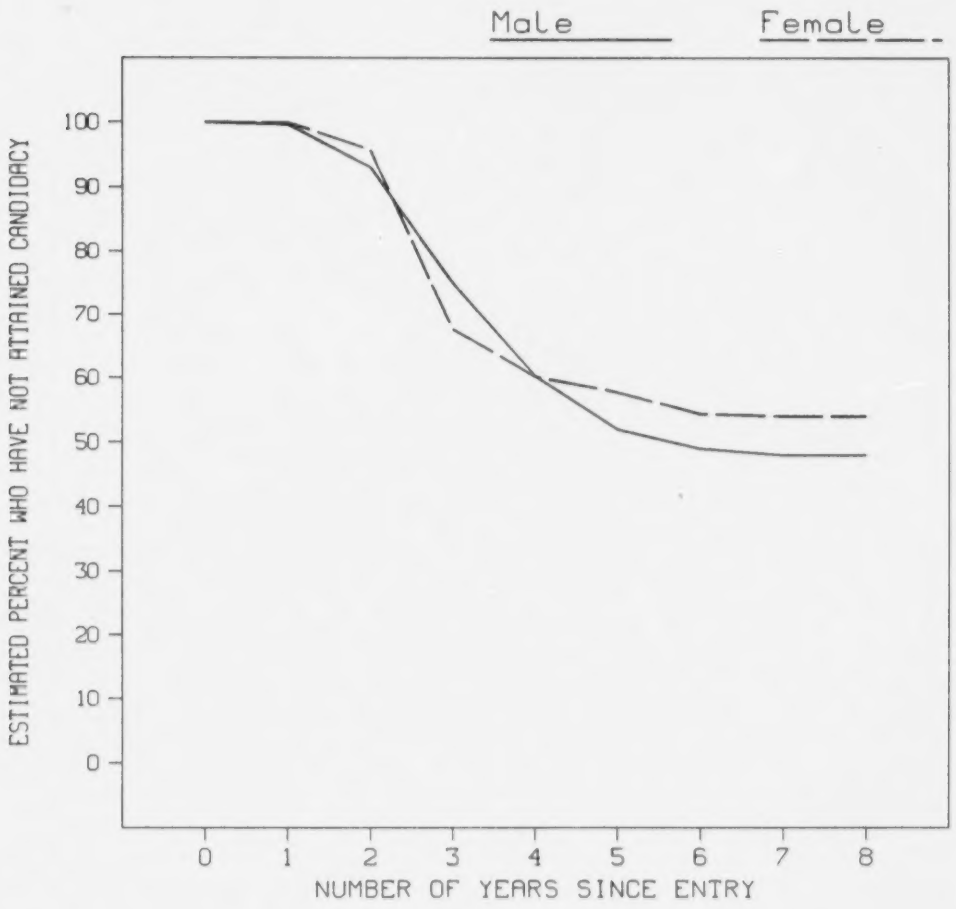
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 1, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 33

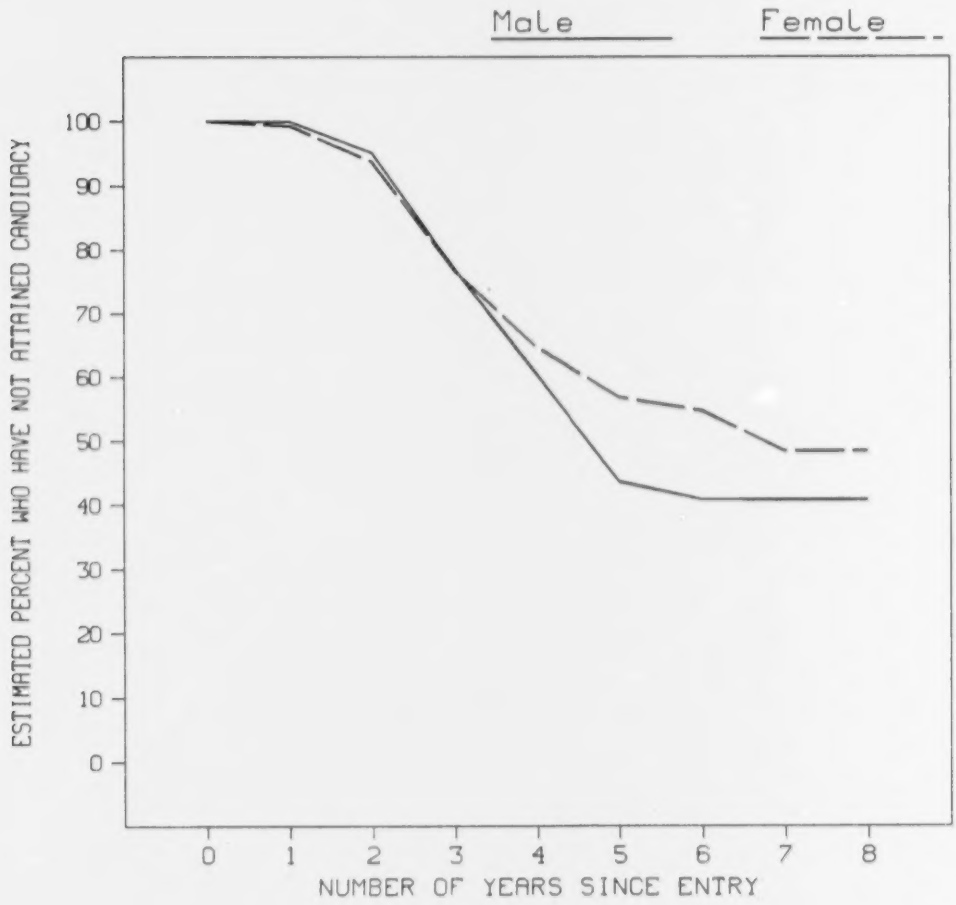
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 2, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 34

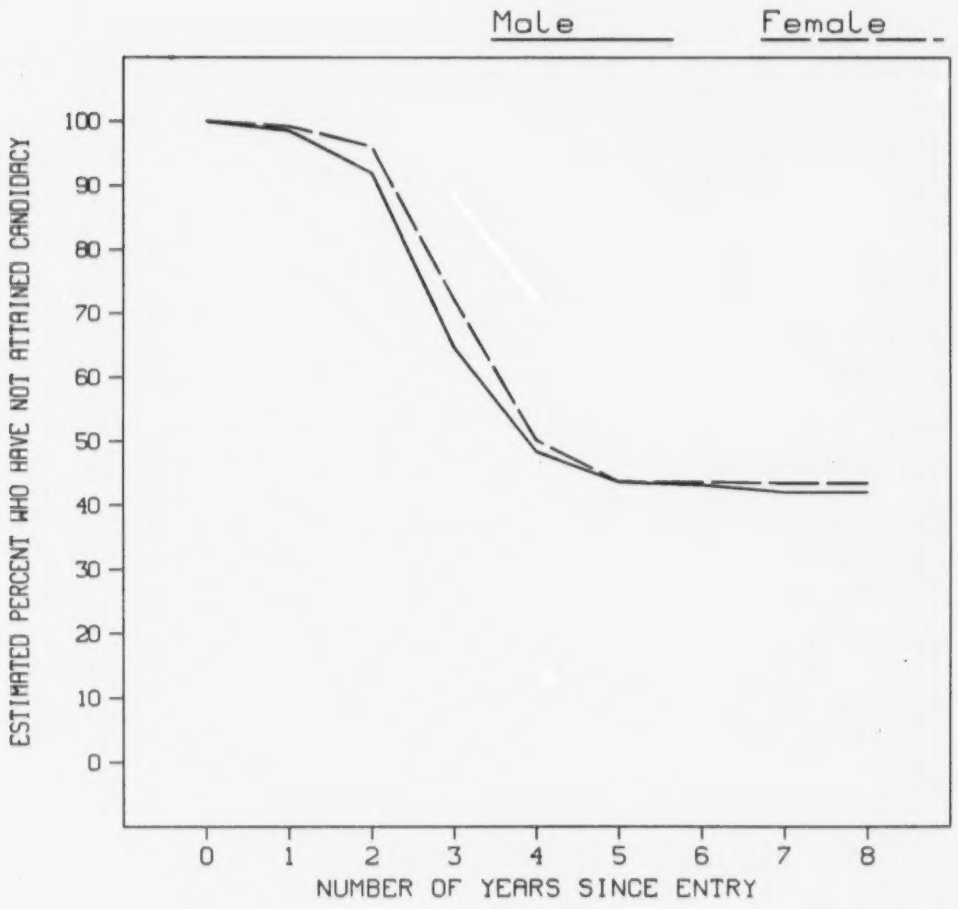
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 2, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 35

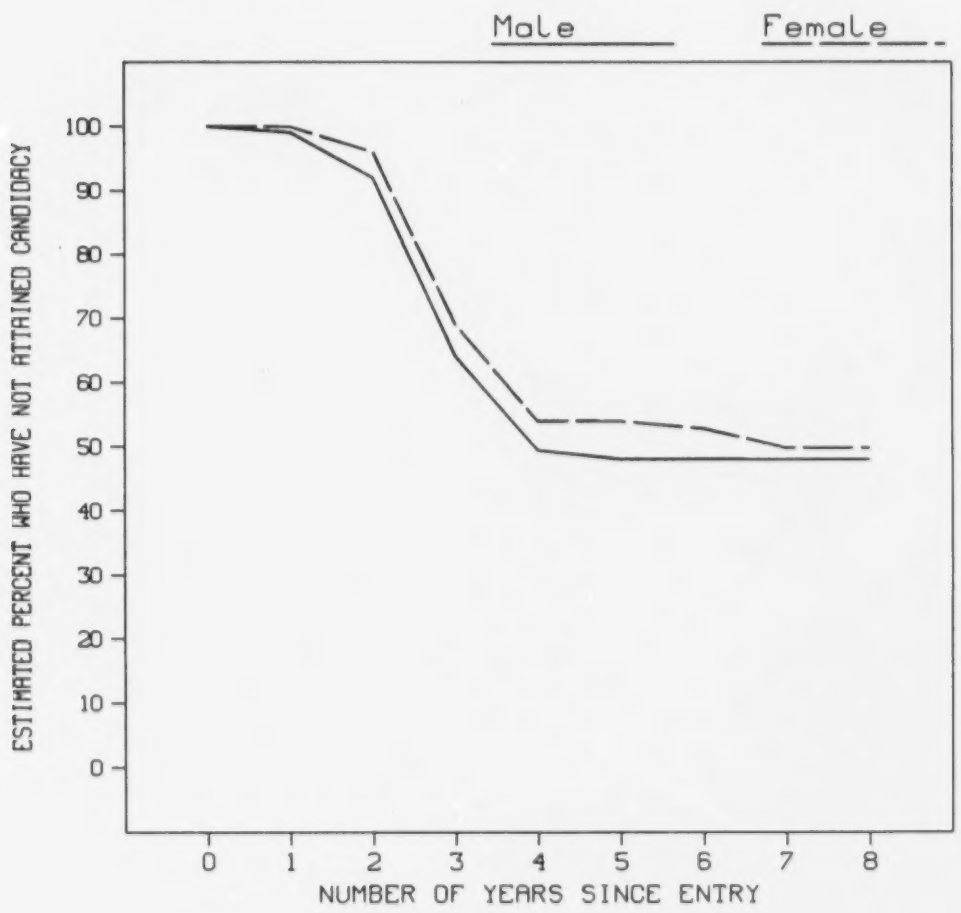
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 3, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 36

BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR CANDIDACY: SCHOOL 3, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

Table 8

Estimated Percentages of Students Achieving Ph.D. Candidacy by Gender Group[a]

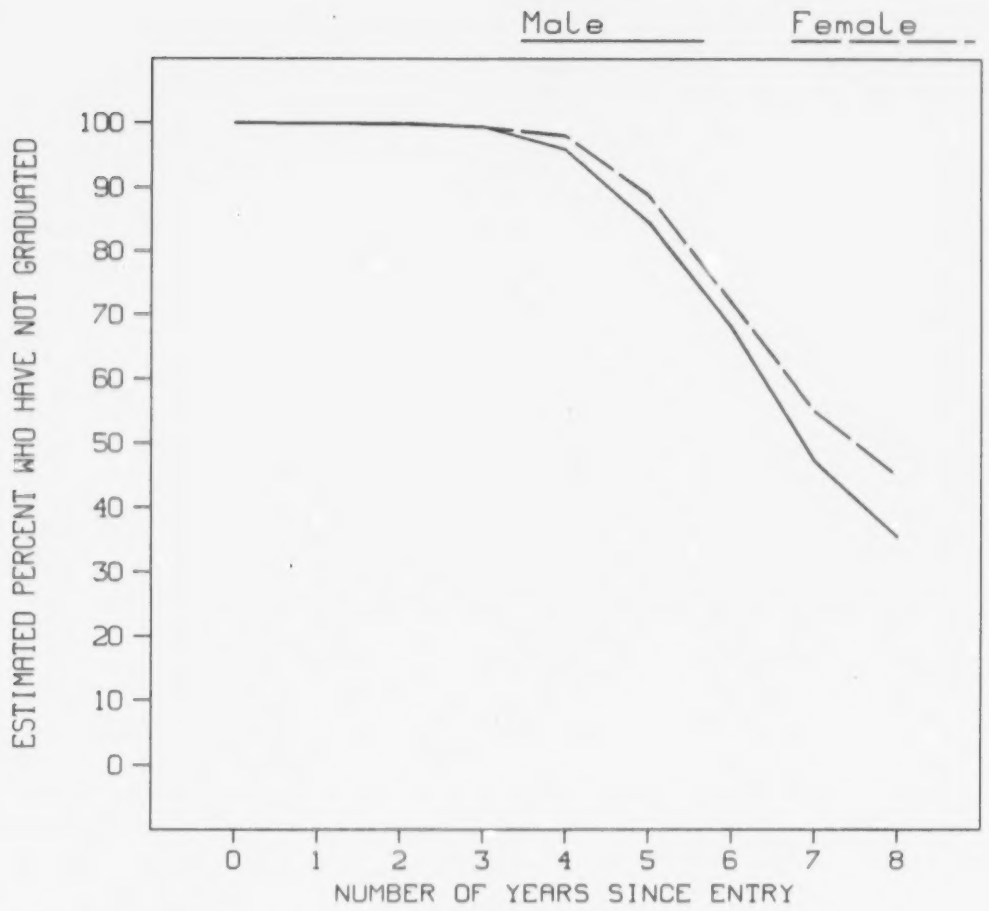
	Five Years After Entry			Eight Years After Entry		
	School 1	School 2	School 3	School 1	School 2	School 3
Cluster 1[b]						
Men	58	48	56	75	52	58
Women	51	42	56	69	46	57
Cluster 2						
Men	53	56	52	64	59	52
Women	45	43	46	59	52	50

[a]The tabled quantities are $100(1 - \hat{S}(x))$, where x is the number of years since entry. Estimates of candidacy and graduation rates are based on slightly different groups of students. See section on survival analysis.

[b]Cluster 1 includes Mathematics, Physics, Computer Science, Economics, and Psychology. Cluster 2 includes Philosophy, History, English, Sociology, and Political Science.

FIGURE 37

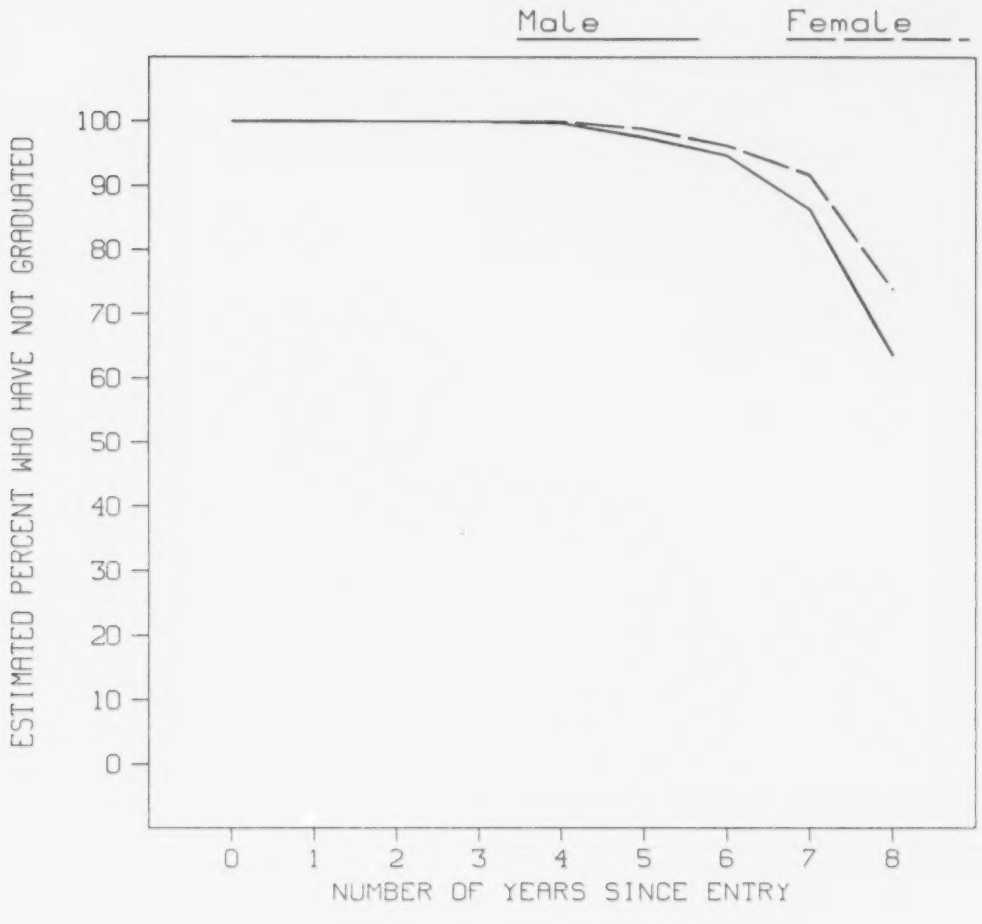
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 1, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 38

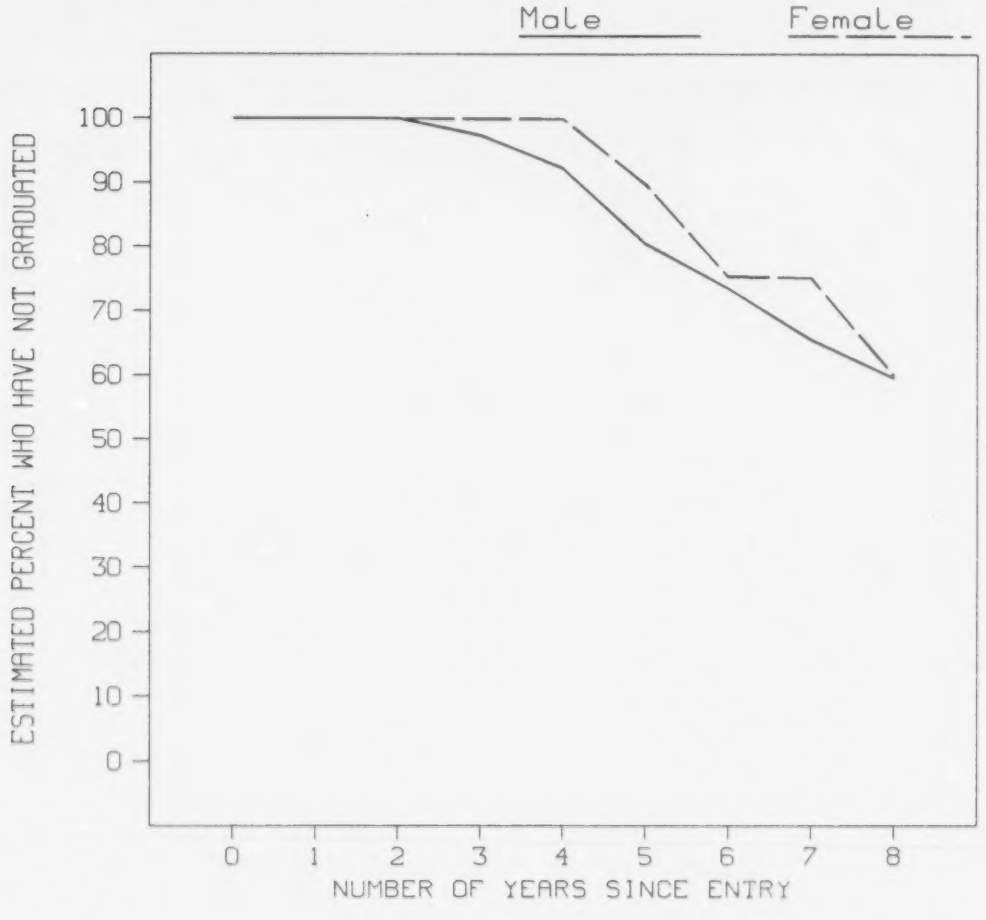
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 1, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 39

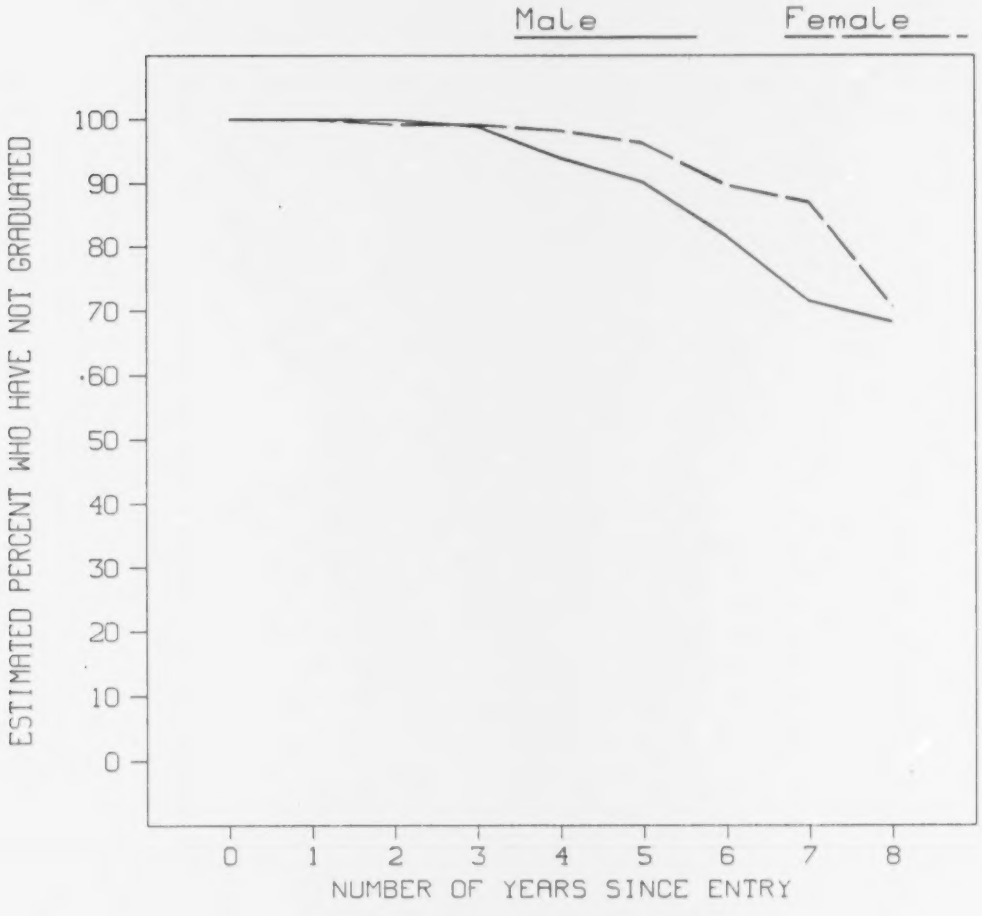
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 2, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 40

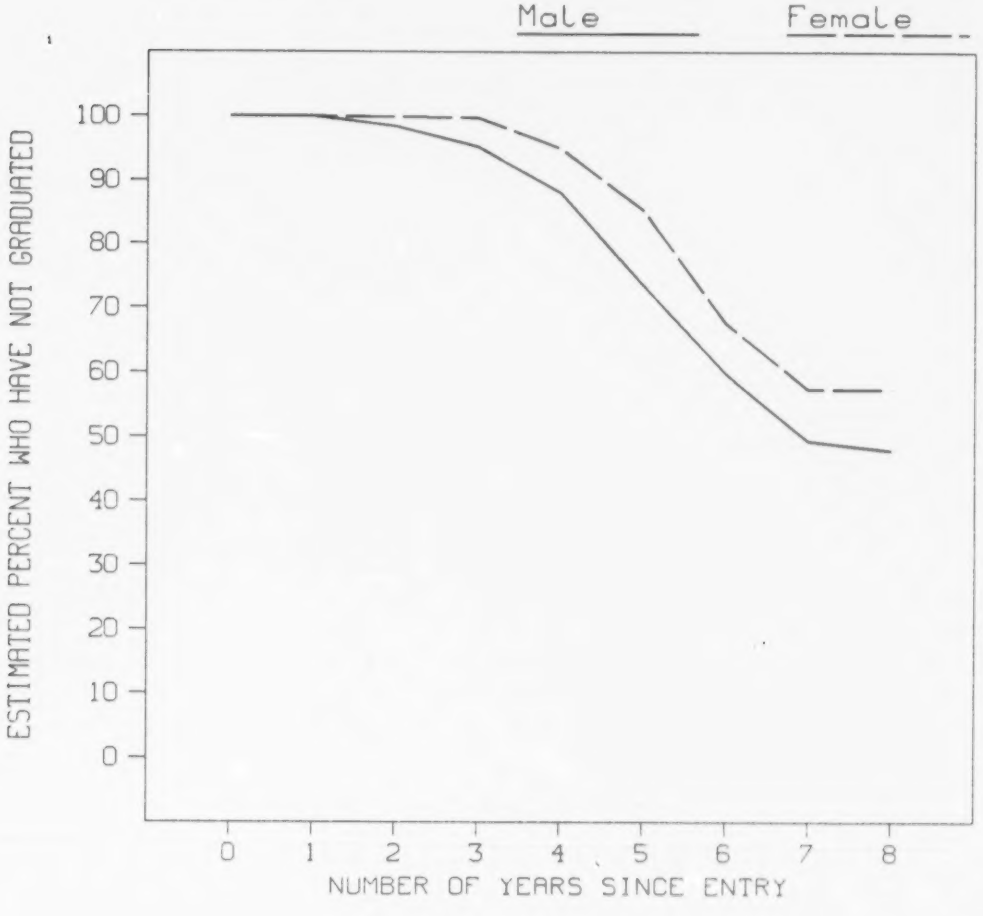
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 2, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 41

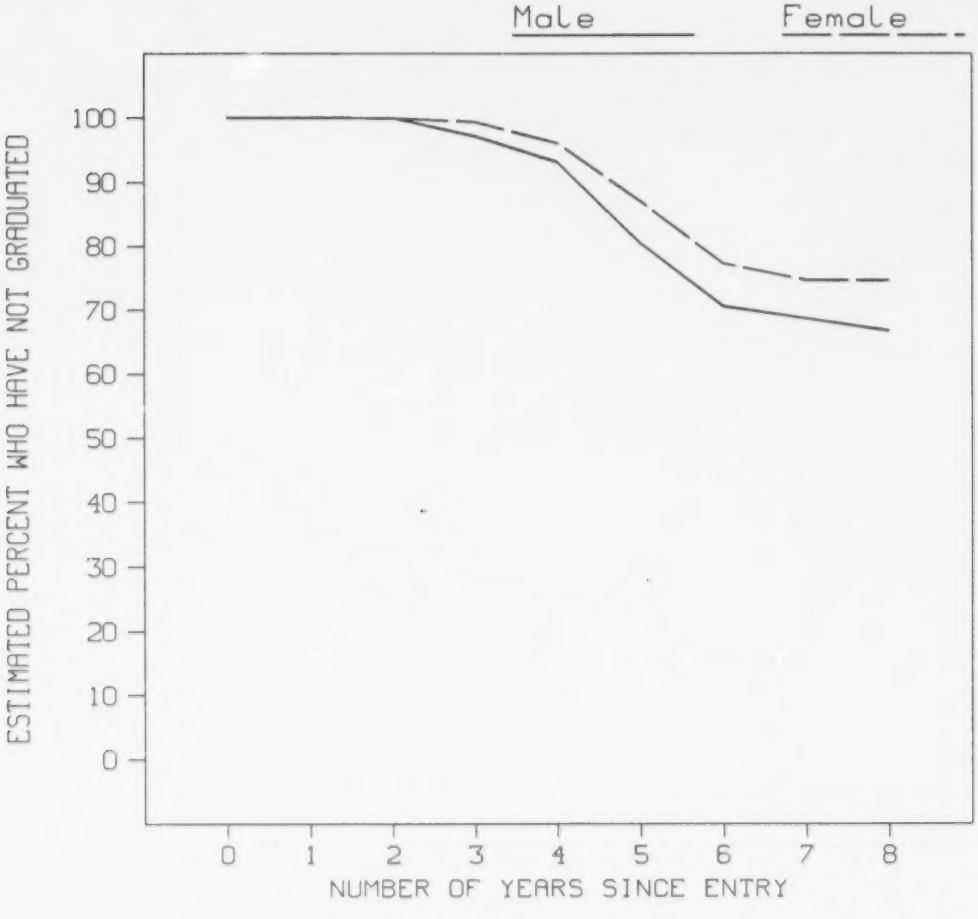
BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 3, CLUSTER 1



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

FIGURE 42

BAYES ESTIMATES OF SURVIVAL FUNCTIONS
FOR GRADUATION: SCHOOL 3, CLUSTER 2



Note: Survival estimates for candidacy and graduation are based on slightly different groups of students. See section on survival analysis.

Table 9

Estimated Percentages of Students Completing Ph.D. Degrees by Gender Group[a]

	Five Years After Entry			Eight Years After Entry		
	School 1	School 2	School 3	School 1	School 2	School 3
Cluster 1[b]						
Men	15	20	26	65	40	52
Women	11	10	15	55	40	43
Cluster 2						
Men	3	10	20	36	32	33
Women	1	4	13	26	29	25

[a]The tabled quantities are $100(1 - \hat{S}(x))$, where x is the number of years since entry. Estimates of candidacy and graduation rates are based on slightly different groups of students. See section on survival analysis.

[b]Cluster 1 includes Mathematics, Physics, Computer Science, Economics, and Psychology. Cluster 2 includes Philosophy, History, English, Sociology, and Political Science.

in Table 11, this analysis indicated that GRE scores and UGPA were almost entirely unrelated to the achievement of candidacy and graduation.[7]

The candidacy variable was defined as follows: Individuals received a code of 1 if they attained candidacy within four years of entry and a code of 0 otherwise. (That is, both dropouts and those who remained in school without attaining candidacy received a code of 0.) Only students who had at least four years of opportunity to receive candidacy before the last recorded candidacy date (see Table 3) were assigned values for this variable. The graduation variable was similarly constructed, except that students received a 1 if they graduated within <u>six</u> years of entry.[8]

The percentage of students for whom GRE scores were available for this analysis ranged from 75 to 100 across the 11 departments at the three schools. For UGPA, the percentage ranged from 44 to 95. Means and standard deviations of GRE scores and UGPA for the 11 departments in the three schools are given in Table 10, along with the percentages of students for whom predictor information was available.

Table 11 gives information on the association between graduation and candidacy on one hand and the three preadmissions measures--UGPA, GREV, and GREQ--on the other. The table shows, for each school and each pair of variables, the 25th percentile, median, and 75th percentile of the distribution of point-biserial correlations for the 11 departments. (A point-biserial correlation is a Pearson correlation between a dichotomous variable and a continuous variable; see, e.g., Hulin, Drasgow, & Parsons, 1983). The sample sizes on which these correlations are based averaged about 170 at School 1, 40 at School 2, and 50 at School 3.

In general, Table 11 shows that prediction was very poor, with median correlations ranging from -.09 to .15. UGPA and GREQ were somewhat more likely to be positively related to candidacy and graduation than GREV. Examination of the results for individual departments showed that the correlations between GREV and graduation that were largest in magnitude tended to be negative. The variations across departments in the size of the correlations did not appear to follow any consistent pattern. In general, prediction was best at School 3, followed by School 2 and then School 1. The pair of variables which exhibited the best evidence of a positive association was UGPA and candidacy, but even here, most correlations did not exceed .20.

---

[7]Although logistic regression analysis (e.g., Hanushek & Jackson, 1977; see Zwick & Braun, 1988, for a possible empirical Bayes strategy) could have been used to model the relation between milestone attainment and the three preadmissions measures, this was not pursued because of the lack of association revealed by these preliminary correlational analyses.

[8]These somewhat arbitrary cut-points were chosen to ensure that adequate numbers of students who did and did not attain the milestone in question were available for estimating the correlation.

Table 10

Means and Standard Deviations of Preadmissions Variables[a]

		School 1			School 2			School 3		
		GREV	GREQ	UGPA	GREV	GREQ	UGPA	GREV	GREQ	UGPA
Chemistry	Mean	612	724	3.69	453	648	3.34	577	703	3.45
	S.D.	101	68	0.36	121	82	0.37	94	67	0.34
	%	97	97	80	99	99	84	91	91	81
Physics	Mean	632	762	3.78	551	699	3.44	553	726	3.26
	S.D.	111	47	0.29	125	66	0.41	132	61	0.37
	%	95	95	73	97	97	80	75	75	47
Mathematics	Mean	597	753	3.68	381	663	3.31	567	735	3.49
	S.D.	142	60	0.47	162	98	0.44	124	56	0.55
	%	86	86	63	100	100	59	92	90	70
Computer Science	Mean	629	769	3.71	499	708	3.49	464	729	3.51
	S.D.	123	39	0.34	166	84	0.27	149	60	0.33
	%	94	94	74	100	100	63	82	82	44
English	Mean	701	574	3.74	629	518	3.38	686	579	3.63
	S.D.	70	108	0.26	102	113	0.42	82	114	0.41
	%	96	96	66	100	100	91	95	95	86
Philosophy	Mean	706	662	3.82	596	620	3.44	682	645	3.60
	S.D.	78	99	0.16	107	107	0.37	64	79	0.33
	%	88	88	54	97	97	95	96	96	88
History	Mean	667	602	3.68	590	558	3.38	651	569	3.64
	S.D.	105	113	0.40	133	89	0.44	95	114	0.30
	%	93	93	69	100	100	93	86	86	91
Psychology	Mean	615	601	3.71	623	643	3.53	631	651	3.57
	S.D.	96	97	0.29	80	76	0.32	101	94	0.33
	%	97	97	75	100	100	93	95	95	85
Political Science	Mean	655	639	3.72	526	527	3.37	544	564	3.44
	S.D.	100	107	0.29	119	99	0.36	131	128	0.34
	%	94	94	71	100	100	91	82	82	69
Sociology	Mean	609	571	3.68	514	597	3.43	558	527	3.45
	S.D.	126	118	0.35	123	116	0.32	106	107	0.39
	%	88	88	60	100	100	82	86	86	76
Economics	Mean	615	715	3.65	536	689	3.44	594	709	3.57
	S.D.	132	78	0.36	128	72	0.37	123	84	0.33
	%	95	95	69	98	98	89	91	91	62

[a]The third entry in each cell is the percentage of students in the department for whom data on the indicated variable were available.

Table 11

Summary Statistics: Correlations Between Milestone Attainment and Pre-admissions Measures for 11 Graduate Programs

	School 1			School 2			School 3		
Correlation	25th %ile	Median	75th %ile	25th %ile	Median	75th %ile	25th %ile	Median	75th %ile
Candidacy[a] with									
UGPA	.02	.08	.17	-.01	.15	.24	-.07	.11	.23
GREV	-.06	-.03	.09	-.19	-.02	.16	-.08	.03	.22
GREQ	.01	.09	.14	-.02	.10	.15	-.13	.09	.31
Graduation[b] with									
UGPA	.01	.04	.12	-.05	0	.27	-.06	.07	.23
GREV	-.19	-.01	.08	-.31	-.09	.05	-.22	-.06	.11
GREQ	-.10	.08	.12	-.10	0	.18	-.03	.11	.20

[a] The candidacy variable is equal to 1 for those who achieved candidacy within four years of entry and equal to 0 otherwise. It is defined only for those with at least four years of opportunity.

[b] The graduation variable is equal to 1 for those who graduated within six years of entry and equal to 0 otherwise. It is defined only for those with at least six years of opportunity.

Correlations of GREV, GREQ, and UGPA with reciprocal time to candidacy and reciprocal time to degree (not shown) also tended to be very low.

Correlational analyses were also conducted within demographic groups. Within each of the three schools, correlations for the six pairs of variables in Table 11 were examined for foreign and White students and for men and women in Clusters 1 and 2. (Sample sizes were too small to permit examination of these correlations for ethnic minority groups.) In the analysis of foreign and White students, most of the 72 correlations (2 demographic groups x 2 clusters x 3 schools x 6 pairs of variables) were close to zero, although there was weak evidence of a positive association between candidacy and UGPA for foreign students in Cluster 1 at School 1 ($r = .26$, $n = 123$) and White students in Cluster 2 at School 3 ($r = .22$, $n = 127$), between candidacy and GREQ for foreign students in Cluster 1 at School 3 ($r = .27$, $n = 77$), and between graduation and UGPA for foreign students in Cluster 1 at School 1 ($r = .37$, $n = 87$).

Prediction was generally very poor in the gender group analysis as well. Again, most of the 72 correlations were near zero. There was some evidence of a positive relation between candidacy and UGPA for men in Cluster 2 in School 2 ($r = .36$, $n = 91$) and women in Cluster 2 in School 3 ($r = .34$, $n = 66$), between candidacy and GREV for women in Cluster 1 in School 2 ($r = .27$, $n = 54$) and women in Cluster 2 in School 3 ($r = .32$, $n = 71$) and between candidacy and GREQ for women in Cluster 2 in School 3 ($r = .31$, $n = 71$).

These results show that, in the populations of Ph.D.-seeking students in these three schools, conventional measures of academic skills could not discriminate well between students who did and did not achieve candidacy and graduation. In evaluating these results, it is important to consider that it was not possible in this study to distinguish among several types of students who did not attain candidacy or graduation by the criterion date: those who were still in school, those who had temporarily left school, those who had withdrawn voluntarily, those who had been asked to withdraw for academic reasons, and those who had been asked to withdraw for nonacademic reasons. It may be, for example, that preadmissions measures could have distinguished between those who completed degrees and those who were asked to withdraw for academic reasons.

In any case, the results do not imply that the GRE and UGPA were useless as admissions criteria: These graduate school matriculants had already been selected on the basis of GRE scores, UGPA, and other factors, and those with the least potential for achieving candidacy or graduation are likely to have been weeded out. From this perspective, then, the low correlations are not unexpected (see Dawes, 1975; Rubin, 1980). Similar results were reported by Zwick and Braun (1988) on the relation of UGPA and GRE scores to graduation and candidacy at Northwestern and by Zwick (1990) on the ability of UGPA and of scores on the Graduate Management Admission Test (GMAT) to predict graduation from doctoral programs in business and management.[9]

---

[9]In a summary of previously conducted studies of the relation between GRE scores and Ph.D. attainment, Willingham (1974) reported median correlations of .18 for GREV and .26 for GREQ. These results are not directly comparable to the present findings because the 47 correlations on which each median was based

There is some evidence that, at the undergraduate level, admissions test results and preadmissions grades also have little association with persistence toward the degree. Willingham (1985) obtained the biserial correlations between a composite of high school rank and SAT and persistence to the senior year of college. These correlations were found to be very low; in six of the nine colleges studied, they did not reach statistical significance.

Within the select population of graduate students, it is likely that such personality factors as perseverance, as well as the availability of financial, social, and faculty support, play a crucial role in determining whether candidacy and graduation are achieved. In a study that included a student survey, Girves and Wemmerus (1988) found that involvement in the graduation program (e.g., participation in research projects, seminars, meetings, and social activities), student relationships with faculty, and financial support had a direct or indirect effect on progress toward the doctoral degree.

As Zwick and Braun (1988) noted in their study of Northwestern University students:

> These findings suggest that further research on candidacy and graduation rates should focus on non-cognitive factors. It may be that improvements in candidacy and graduation rates can best be achieved by designing admissions procedures that place more weight on personality attributes like determination or persistence and by improving support systems for students already in school. (p. 38)

---

came from different institutions and corresponded to different administrative units.

Summary and Discussion

Several types of analyses were conducted, based on nearly 5,000 Ph.D.-seeking students who matriculated in 11 departments at each of three large research universities between 1978 and 1985.

Descriptive analyses (combining across the 11 departments) showed that at all three schools, 72 to 75% of the students who entered between the fall of 1978 and the summer of 1981 were White and 18 to 20% were foreign. The percentages of students who were Asian Americans, Black Americans, or Hispanic Americans summed to only 5 to 7%. For students who entered between fall of 1981 and fall of 1985, results at School 1 stayed essentially the same, while at the other two schools, the percentage of foreign students increased by about 10 and the percentage of Whites decreased by the same amount. Results for the individual departments showed that most departments at School 2 and School 3, as well as the Computer Science and Sociology departments at School 1, experienced an increase in the percentage of foreign students and a decrease in the percentage of Whites.

Examination of the proportions of men and women in graduate school showed that men outnumbered women by about 2 to 1 or more in each cohort at each of the three schools, combining across the 11 departments. At all three schools, the percentage of men was between 72 and 74 for students who entered between fall of 1978 and summer of 1981. (A Carnegie Commission study in 1968 yielded similar results for a sample of 80 universities [Feldman, 1974, p. 15].) For students who entered between fall of 1981 and fall of 1985, the percentages remained relatively steady at School 1 and School 3 but dropped to 65% at School 2. The percentage of men at all three schools was highest for foreign students, followed in order by Asian, White, Hispanic, and Black students.

Considering the data from all three schools, the most heavily male departments were Mathematics, Physics, and Computer Science, all of which were typically at least 80% male. The departments that came closest to having equal numbers of men and women were English, History, Psychology, and Sociology.

Survival analysis methods were used to study the rates of candidacy and graduation for the eight years following students' entry into graduate school. First, analyses were conducted by department. For purposes of display and discussion, departments were divided into three groups. In general, the Group I departments--Chemistry, Physics, Mathematics, and Computer Science--had higher rates of candidacy and graduation than did the Group III departments--Psychology, Political Science, Sociology, and Economics--which, in turn, had higher attainment rates than did the Group II departments--English, Philosophy, and History.

Only at School 1 were at least 50% of students in all 11 departments estimated to have achieved candidacy by year 8. For several departments,

survival functions were still decreasing at year 8. In about half the departments at School 1 and most departments at the two other schools, the estimated percentage of students graduating by year 8 was less than 50.

Survival analyses were also conducted for ethnic and gender groups. A finding that was generally consistent for Schools 1 and 3 was that candidacy and graduation rates were higher for foreign students than for White Americans and higher for White Americans than for Black Americans. Results from School 1 showed that Asian and Hispanic Americans also tended to have lower attainment rates than White students.

Because survival analysis allows examination of candidacy and graduation rates at multiple time points, it gives a more detailed picture of milestone attainment than do simple rates of candidacy or graduation. For example, Table 5 allows us to state that, of an entering class of 10 students in English at any of the three schools, only 2 to 3 students would be expected to receive the doctorate by eight years after entry. Analyses of this type may be useful to graduate school deans in estimating the number of graduates an entering class is likely to yield and in determining whether administrative changes are needed to hasten progress in some departments or whether special programs are needed to encourage the persistence of some groups of students.

The well-documented scarcity of Black Americans and other minorities enrolling in U.S. graduate schools (ACE, 1987; ACE/ECS, 1988; Blackwell, 1987; Brown, 1987; Mooney, 1989; Trent & Copeland, 1987) was strikingly evident in the three schools in this study. Also, the candidacy and graduation rates of Asian, Black, and Hispanic Americans were lower than those of White Americans and foreign students. What accounts for the low participation rates of minority students in graduate education? One of the most commonly mentioned possible reasons is the lack of adequate financial resources (Blackwell, 1987; Mooney, 1989; National Board of Graduate Education, 1976; Nettles, 1987; Pruitt & Isaac, 1985; Thomas, 1987). Another possible factor is the attraction of professional schools (Chamberlain, 1988; Mooney, 1989; Pruitt & Isaac, 1985), although Nettles (1987) and Thomas (1987) concluded that this speculation was unsupported. Some researchers have cited discrimination in the recruitment and admissions process (Pruitt & Isaac, 1985), the perceived "inhospitality of academe" (Mooney, 1989), and the lack of adequate support services and opportunities for faculty mentoring for minorities. Finally, some sources (Astin, 1982; Blackwell, 1987; National Board of Graduate Education, 1976) have stressed the need for increased support for minority students earlier in the educational pipeline. A 1988 ACE/ECS report states:

> The aptitude for higher education and the ability to succeed in college and graduate school do not materialize suddenly at age 18; they are developed in childhood. Currently, we lose disproportionate numbers of minority students at each level of schooling, culminating in low participation rates in higher education. Only through intense, coordinated efforts at every stage—beginning with adequate prenatal care, improved nutrition,

> and quality child care and extending through programs to increase minority retention and improve student performance at the elementary and secondary levels—can we hope to reverse these dismal trends. (p. 14)

Another finding of the current study was that, at Schools 1 and 3, foreign students had higher candidacy and graduation rates than did American students. There are several possible reasons for these higher rates. Foreign students are likely to have been selected to study in the United States because of their academic excellence. Girves and Wemmerus (1988, p. 169) speculated that "the fact that foreign students must be enrolled full-time and must demonstrate sufficient financial support to carry out their degree programs may be more incentive for them to complete their degrees. Domestic students, on the other hand, do not necessarily have these incentives, and may have other options outside of graduate school." It is not clear why the pattern at School 2, where Whites had rates that equaled or exceeded those of foreign students, differed from that observed at School 1 and School 3.

In general, both candidacy and graduation rates tended to be higher for men than for women. Earlier studies have also found that women enrolled in doctoral programs are less likely than their male counterparts to attain the degree (see Feldman, 1974; Patterson & Sells, 1973). Some of the reasons cited for the lower proportions of women in graduate school and their lower attainment rates are analogous to those offered in discussing the participation of ethnic minorities in graduate school--for example, discrimination in admissions, in counseling, (Roby, 1973), and in the distribution of financial aid (Chamberlain, 1988), and an inhospitable campus climate (Chamberlain, 1988; Feldman, 1974; Schwartz & Lever, 1973). Berg and Ferber (1983) and Feldman (1974) cited the scarcity of female mentors in graduate school, along with the finding that women were less likely than men to establish close working relationships with male faculty members. In discussing the higher dropout rates for women in graduate programs in science and engineering, Widnall (1988) cited lack of self-esteem, feelings of alienation, poor relationships with faculty advisors, and discrimination as possible reasons. Women may also be less likely than men to perceive the completion of school as an economic necessity, as Tinto (1975) noted with regard to undergraduate education. It may also be true that women are more likely than men to take reduced course loads or to leave school because of family responsibilities.

Examination of GRE and UGPA data for ethnic and gender groups showed that group attainment rates were not, in general, ordered in the same way as group means on the preadmissions measures. For example, at School 1 in both clusters, White students had roughly the same GREQ and UGPA means as foreign students and had GREV means that were about 150 points higher, yet foreign students had higher attainment rates. Similarly, Hispanics in School 1 tended to have higher means than Black students on the preadmissions measures, but had lower attainment rates.

Finally, more formal and comprehensive analyses were conducted to examine the relation between candidacy and graduation on one hand and pre-admissions measures of academic skills on the other. These correlational

analyses also showed little or no relation between attainment rates and UGPA or GRE scores. This finding is not entirely unexpected, given that graduate students have been selected using these measures. Within this very competent group of people, it is likely that nonacademic measures determine who will succeed.

It is important to note that all three universities in this study have selective and prestigious graduate schools. Therefore, the research results cannot be assumed to be widely generalizable. However, it is hoped that the findings will be useful to the participating schools, will serve to illustrate the types of analyses that can be informative to graduate school policy makers, and will suggest hypotheses that can be investigated for larger groups of institutions.

APPENDIX A

Letter of Request for Participation

Dear _____:

The Graduate Record Examinations Board and Educational Testing Service are sponsoring a new study, "Analysis of Pathways Through Graduate School for Minority Students." This research project will involve examination of patterns of attainment of candidacy and doctoral degrees, focusing on the graduate careers of minorities. The association between the successful completion of graduate school requirements and measures of academic performance, such as undergraduate grade-point average and GRE scores, will also be investigated. Universities with centralized computer data bases on graduate students are being sought for inclusion in the study. Participating universities may elect to remain anonymous in reports of the research findings. Further detail on the study, as well as a description of the data required, is given in the enclosed abstract.

I hope that it will be possible for [your university] to participate. I will call you within the next several weeks so that we can discuss the project further. If you have any questions, please feel free to call me at 609-734-5311.

Sincerely,

Rebecca Zwick, Ph.D.
Project Director

Enc.

APPENDIX B

Description of Project

Analysis of Pathways Through Graduate School for Minority Students

Rebecca Zwick, Project Director
Educational Testing Service
July, 1987

Many observers believe that graduate education in the United States is at a critical juncture. A survey issued by GREB/CGS in 1985 shows decreases in first-time graduate enrollment in all disciplines. Furthermore, these decreases are not uniform across ethnic groups. For example, between 1983 and 1984, first-time graduate enrollment decreased 7.5% for Blacks, while the decrease for Whites was only 1.4%. Graduate school deans are now faced with the challenge of analyzing these trends and developing appropriate policies concerning admissions, retention and academic standards.

It is important, therefore, to determine what happens to those individuals who actually enroll in graduate school. At what pace do these students reach milestones in their graduate careers, such as advancement to candidacy and attainment of the doctoral degree? What attributes differentiate students who complete the doctorate from those who do not? How do the patterns of achievement differ across academic programs and across ethnic groups?

ETS and the GRE Board are jointly sponsoring a multi-institutional study to investigate these issues. The study is a part of the GREB Minority Graduate Education Project, which is linked to efforts to understand and increase minority participation in graduate education that are being undertaken by the CGS and the AGS. The research questions and analyses fall into two broad categories:

1. How do the patterns of timing and attainment of graduate school milestones, such as candidacy and graduation, differ across academic disciplines and across ethnic groups? Analyses that will address this question will include simple descriptive statistics, such as proportions of matriculants in various disciplines and ethnic groups who completed each milestone. In addition, the data provided in graduate school term records lend themselves well to survival analysis, a statistical method often applied in clinical trials. In the present context, survival analysis will involve estimation of the probability that a graduate school milestone, such as candidacy or degree completion, will take more than a specified number of years to occur. Separate survival curves can be plotted for each academic discipline and for demographic groups of interest.

2. What is the association between students' completion of candidacy and degree requirements and measures of their academic performance, such as undergraduate grade-point average and Graduate Record Examination scores? How do the patterns of association differ across academic disciplines and ethnic groups? The analyses that will address these questions include simple correlations, as well as logistic regression analysis, which is based on a model for the probability that an event occurs.

In investigating minority issues, it is desirable to examine separately the graduate career of each minority group of interest. A familiar complication in this type of research is the small sample sizes which are often encountered. The statistical problems associated with small sample size are reduced to some degree by the use of empirical Bayes methods which have been applied successfully in related ETS studies.

University Participation

We are now contacting universities to request their participation. We are seeking schools with centralized computerized data bases on their doctoral students, preferably dating back to 1980 or before. The student data of primary interest are:

- Entry date and department
- Ethnicity and citizenship
- Gender
- Undergraduate grade-point average
- GRE verbal score
- GRE quantitative score
- Date of advancement to candidacy
- Date of graduation
- Graduate school grade-point average (possible at completion of first year or at completion of course work)

We would require a tape of the data, as well as documentation of the tape position and definition of the included variables. In addition, we would need the name of a contact person who could answer questions about the data. We will be able to reimburse universities for costs involved in data preparation. Participating universities may elect to remain anonymous in reports of the research findings. We hope to collect all data by the fall of 1987 and to issue a report of the results by the fall of 1989. We believe that this study will benefit the graduate community and that the results will be particularly useful to participating institutions.

APPENDIX C

Description of Data Requirements

August 21, 1987

[address]

Dear :

I am very pleased that [name of university] has agreed to participate in the study, "Analysis of Pathways Through Graduate School for Minority Students," which is jointly sponsored by the Graduate Record Examinations Board and Educational Testing Service. So far, four universities (three public and one private) have agreed to participate and several others have expressed an interest in the project.

The study will involve the examination of the patterns of timing and attainment of candidacy and graduation and the relation between the attainment of these milestones and measures of academic potential. The focus will be the graduate careers of minority students. The GREB and ETS believe that it is important to conduct a multi-institutional study in order to obtain a broad picture of the graduate school experiences of minorities. The study is part of the GREB Minority Graduate Education Project, which is linked to efforts to understand and increase minority participation in graduate education that are being undertaken by the CGS and the AGS. The final report of the results, which is expected to be completed by the fall of 1989, will be provided to all participating institutions.

Participating universities may elect to remain anonymous in reports of the research findings. Please let me know, in writing, whether the identity of your university can be disclosed in reports and publications and whether other participating universities may be informed that your university is taking part in this research.

The student population of interest consists of graduate students (both minorities and whites) who are seeking doctorates. The student data of primary interest are:

Entry date
Department or graduate program
Ethnicity (at least for U.S. citizens)
Citizenship (i.e., whether student is a U.S. citizen)
Gender
Undergraduate grade-point average
GRE verbal score
GRE quantitative score
Date of advancement to candidacy
Date of graduation
Date or term of departure from school if degree not completed

I am aware that it will not be possible to obtain all these data from every participating institution. Your university's data will be valuable to the project even if only a portion of this information is available. [*confirm type of data available.]

If possible, the data tape should be a 9-track IBM standard label tape that uses an EBCDIC character set and has a density of 6250 bpi. If it is not possible to produce a tape with these specifications, other kinds of tapes can also be accommodated. I will also need documentation that includes the file position of all variables, the meanings of all codes, including missing data codes, and information about any unusual features of the data. In addition, if a general catalog of graduate programs is available, I would appreciate it if you would send me a copy. I am hoping to receive all data tapes and documentation by October, 1987. Participating universities may be reimbursed for costs associated with the preparation and mailing of tapes by sending me an itemized bill.

Once again, I am delighted to hear that [*university] [*has agreed to participate/is considering participation] in this study. I truly appreciate your willingness to undertake the work involved in providing the necessary data. If you have any questions about the project, please feel free to call me at 609-734-5311.

Sincerely,

Rebecca Zwick, Ph.D
Project Director

APPENDIX D

Supplementary Tables

Table D-1

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: Chemistry

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	223	425	37	68	77	138
Ethnic Group						
Asian	5	4	0	1	3	2
Black	1	0	0	0	0	1
Hispanic	1	1	0	0	3	1
White	85	82	68	68	92	78
Other	2	1	0	1	0	0
Foreign	4	11	27	29	3	19
Missing	0	0	5	0	0	0
Gender Group						
Male	83	77	89	68	70	75
Female	17	23	11	32	30	25

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-2

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: English

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	81	172	45	76	25	34
Ethnic Group						
Asian	6	2	0	3	0	3
Black	0	2	4	1	8	0
Hispanic	0	2	2	0	0	3
White	83	88	82	84	76	85
Other	2	0	0	1	0	0
Foreign	9	6	11	7	12	9
Missing	0	0	0	4	4	0
Gender Group						
Male	40	51	49	51	52	44
Female	60	49	51	49	48	56

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-3

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: History

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	94	170	10	17	24	42
Ethnic Group						
Asian	1	2	0	0	0	2
Black	4	1	10	0	21	7
Hispanic	2	4	0	0	0	0
White	82	78	90	71	71	86
Other	0	2	0	0	0	0
Foreign	11	12	0	29	8	5
Missing	0	2	0	0	0	0
Gender Group						
Male	62	64	70	41	50	60
Female	38	36	30	59	50	40

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-4

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: Mathematics

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	126	233	11	48	21	40
Ethnic Group						
Asian	2	2	0	4	0	0
Black	2	1	0	0	0	2
Hispanic	0	4	9	2	0	0
White	55	52	27	17	90	67
Other	4	1	0	0	0	0
Foreign	37	38	64	75	5	30
Missing	1	1	0	2	5	0
Gender Group						
Male	90	81	82	77	86	88
Female	10	19	18	23	14	13

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-5

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: Political Science

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	83	150	13	10	37	43
Ethnic Group						
Asian	4	3	8	0	0	2
Black	1	0	15	0	11	9
Hispanic	2	4	0	0	0	0
White	64	75	54	90	57	49
Other	2	1	0	0	0	0
Foreign	25	16	23	10	32	40
Missing	1	0	0	0	0	0
Gender Group						
Male	71	65	62	60	78	72
Female	29	35	38	40	22	28

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-6

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: Psychology

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	67	102	35	52	25	41
Ethnic Group						
Asian	4	6	3	4	0	0
Black	7	3	0	2	0	5
Hispanic	3	4	3	4	4	2
White	70	74	83	79	92	78
Other	4	5	0	0	0	0
Foreign	9	9	11	12	4	15
Missing	1	0	0	0	0	0
Gender Group						
Male	42	35	63	42	56	61
Female	58	65	37	58	44	39

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-7

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: Economics

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	86	127	29	64	77	106
Ethnic Group						
Asian	1	3	0	0	1	2
Black	2	2	0	0	0	2
Hispanic	1	2	0	0	0	1
White	64	65	76	52	69	57
Other	1	0	0	0	0	0
Foreign	30	28	21	45	29	39
Missing	0	0	3	3	1	0
Gender Group						
Male	73	66	83	81	74	81
Female	27	34	17	19	26	19

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-8

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: Philosophy

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	20	32	13	24	26	25
Ethnic Group						
Asian	0	6	0	0	0	4
Black	0	0	0	0	0	0
Hispanic	0	0	0	0	0	0
White	55	72	100	79	100	92
Other	0	0	0	0	0	0
Foreign	45	19	0	21	0	4
Missing	0	3	0	0	0	0
Gender Group						
Male	75	63	85	71	85	84
Female	25	38	15	29	15	16

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-9

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: Physics

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	139	184	7	28	36	65
Ethnic Group						
Asian	3	3	0	0	0	2
Black	1	0	0	0	3	0
Hispanic	1	1	0	0	0	0
White	73	73	86	68	53	42
Other	1	1	0	0	0	0
Foreign	20	21	14	29	44	57
Missing	2	2	0	4	0	0
Gender Group						
Male	91	92	100	93	89	85
Female	9	8	0	7	11	15

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-10

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: Computer Science

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	78	182	6	18	42	93
Ethnic Group						
Asian	4	5	0	0	0	2
Black	0	0	0	0	0	1
Hispanic	1	2	0	0	0	0
White	68	55	67	28	50	35
Other	4	1	0	0	0	0
Foreign	23	36	33	72	48	61
Missing	0	2	0	0	2	0
Gender Group						
Male	92	89	83	89	98	82
Female	8	11	17	11	2	18

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-11

Percentages[a] of Ethnic and Gender Groups
for Two Cohorts[b] in Three Schools: Sociology

	School 1		School 2		School 3	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Sample Size	61	78	4	18	33	41
Ethnic Group						
Asian	3	4	0	0	0	0
Black	5	9	0	0	6	5
Hispanic	7	6	0	0	9	5
White	67	54	75	67	64	76
Other	3	3	0	0	0	0
Foreign	15	24	25	28	21	15
Missing	0	0	0	6	0	0
Gender Group						
Male	49	62	75	50	42	39
Female	51	38	25	50	58	61

[a]Percentages may not add to 100 because of rounding.

[b]Cohort 1 includes entry dates of fall 1978 through summer 1981. Cohort 2 includes entry dates of fall 1981 through fall 1985.

Table D-12

Sizes of Risk Sets for Survival Analysis of Candidacy for Eleven Departments

Years Since Candidacy	Chemistry	Physics	Mathematics	Computer Science	English	Philosophy	History	Psychology	Political Science	Sociology	Economics
School 1											
1	648	323	359	260	253	52	264	169	233	139	213
2	647	323	354	260	253	52	264	168	233	139	213
3	324	323	304	256	252	51	259	167	233	135	212
4	168	269	173	176	204	35	196	115	191	116	175
5	73	178	122	98	128	25	117	63	106	87	123
6	62	93	90	56	80	19	79	38	60	60	77
7	48	53	69	39	60	17	61	21	46	39	43
8	38	30	48	21	39	10	41	16	33	24	31
School 2											
1	105	35	59	24	121	37	27	87	23	22	93
2	105	34	59	24	120	37	27	87	23	22	93
3	54	30	46	24	117	37	22	86	21	20	93
4	39	19	26	11	85	32	9	39	12	12	79
5	35	9	15	11	58	23	6	26	11	8	54
6	26	5	9	9	35	12	3	20	10	5	37
7	18	4	6	6	24	10	2	16	9	3	24
8	12	4	4	2	14	8	2	12	8	1	15
School 3											
1	215	101	61	135	59	51	66	66	80	74	183
2	215	101	56	135	59	51	66	66	80	72	181
3	44	75	41	98	45	41	59	55	69	59	157
4	27	46	17	52	23	27	36	34	40	36	97
5	23	27	12	38	18	22	18	15	27	21	62
6	17	23	11	26	16	22	13	11	24	18	47
7	11	19	10	24	15	18	9	7	19	15	36
8	7	11	6	16	10	14	7	7	17	8	21

Table D-13

Sizes of Risk Sets for Survival Analysis of Graduation for Eleven Departments

Years Since Candidacy	Chemistry	Physics	Mathematics	Computer Science	English	Philosophy	History	Psychology	Political Science	Sociology	Economics
School 1											
1	648	323	359	260	253	52	264	169	233	139	213
2	648	323	358	260	253	52	264	169	233	139	213
3	648	323	357	260	253	52	264	169	233	139	213
4	568	281	312	208	209	47	234	143	213	128	184
5	462	241	262	165	175	43	197	120	174	112	158
6	271	194	184	108	144	36	145	89	132	88	130
7	114	135	129	77	112	29	122	63	106	73	91
8	54	65	75	35	78	16	84	34	76	52	51
School 2											
1	105	35	59	24	121	37	27	87	23	22	93
2	105	35	59	24	121	37	27	87	23	22	93
3	98	30	50	17	108	34	22	81	21	18	80
4	83	25	38	12	98	27	18	74	20	14	66
5	56	10	28	10	77	25	16	58	14	11	52
6	36	8	13	7	57	17	11	38	13	7	41
7	18	5	6	6	38	13	10	21	10	2	26
8	11	4	3	3	23	8	5	9	7	2	18
School 3											
1	215	101	61	135	59	51	66	66	80	74	183
2	186	83	50	108	51	42	61	59	71	67	161
3	159	69	39	74	37	37	53	49	62	54	142
4	131	55	30	57	31	31	39	41	53	42	123
5	95	45	25	43	26	29	30	31	42	37	92
6	34	33	17	30	21	22	22	19	30	26	63
7	15	18	8	20	15	17	13	12	25	13	34
8	8	8	3	13	7	10	8	7	16	7	16

Table D-14

Sizes of Risk Sets for Survival Analysis of Candidacy for Ethnic Groups

Years Since Entry	Cluster 1					Cluster 2				
	Asian	Black	Hispanic	White	Foreign	Asian	Black	Hispanic	White	Foreign
School 1										
1	43	19	25	839	361	28	20	29	710	136
2	43	19	25	836	358	28	20	29	710	136
3	42	19	25	813	330	28	20	29	705	130
4	23	16	20	607	219	26	17	26	559	98
5	15	12	10	400	129	17	13	21	346	56
6	10	8	9	240	73	15	9	13	218	38
7	5	5	4	151	48	11	7	11	159	30
8	3	3	3	95	34	5	5	6	108	19
School 2										
1				170	112				185	30
2				169	112				184	30
3				165	98				172	30
4				106	54				120	18
5				74	32				83	12
6				49	23				49	8
7				35	16				34	7
8				23	10				23	5
School 3										
1		7		314	213		22		244	53
2		7		311	209		22		243	52
3		4		264	148		22		198	44
4		2		163	75		19		115	22
5		1		98	50		16		72	13
6		1		82	31		11		66	13
7		1		68	23		8		57	10
8		0		47	13		5		42	9

Table D-15

Sizes of Risk Sets for Survival Analysis of Graduation for Ethnic Groups

Years Since Entry	Cluster 1					Cluster 2				
	Asian	Black	Hispanic	White	Foreign	Asian	Black	Hispanic	White	Foreign
School 1										
1	43	19	25	839	361	28	20	29	710	136
2	43	19	25	839	360	28	20	29	710	136
3	43	19	25	839	359	28	20	29	710	136
4	29	18	21	721	309	26	18	26	621	123
5	24	16	14	615	251	22	15	24	524	103
6	20	13	12	468	169	19	12	18	402	85
7	17	9	6	342	104	17	11	15	320	70
8	8	7	4	169	58	11	8	8	226	47
School 2										
1				170	112				185	30
2				170	112				185	30
3				155	88				165	24
4				136	67				145	18
5				101	48				116	16
6				74	27				86	10
7				45	15				59	8
8				26	8				36	4
School 3										
1		7		314	213		22		244	53
2		5		277	169		22		212	48
3		3		237	125		22		171	41
4		3		198	98		19		137	33
5		2		160	68		15		117	27
6		1		116	40		13		88	18
7		0		70	21		9		59	14
8		0		39	7		2		39	7

Table D-16

Sizes of Risk Sets for Survival Analysis of Candidacy for Gender Groups

Years Since Entry	Cluster 1 Male	Cluster 1 Female	Cluster 2 Male	Cluster 2 Female
School 1				
1	1044	280	557	384
2	1039	279	557	384
3	987	275	547	383
4	711	197	434	308
5	464	120	256	207
6	276	78	160	138
7	179	46	120	103
8	113	33	78	69
School 2				
1	220	78	129	101
2	219	78	129	100
3	204	75	123	94
4	131	43	83	67
5	87	28	59	47
6	62	18	33	32
7	44	12	23	25
8	31	6	18	15
School 3				
1	439	107	198	132
2	433	106	196	132
3	342	84	160	113
4	195	51	96	66
5	126	28	62	44
6	98	20	53	40
7	79	17	45	31
8	49	12	35	21

Table D-17

Sizes of Risk Sets for Survival Analysis of Graduation for Gender Groups

Years Since Entry	Cluster 1 Male	Cluster 1 Female	Cluster 2 Male	Cluster 2 Female
School 1				
1	1044	280	557	384
2	1043	280	557	384
3	1042	280	557	384
4	897	231	492	339
5	756	190	409	292
6	559	146	316	229
7	401	94	252	190
8	208	52	171	135
School 2				
1	220	78	129	101
2	220	78	129	101
3	188	70	114	89
4	158	57	100	77
5	115	43	80	63
6	79	28	60	45
7	50	14	41	32
8	31	6	25	20
School 3				
1	439	107	198	132
2	374	87	175	117
3	303	70	147	96
4	248	58	115	81
5	188	48	99	65
6	127	35	74	47
7	74	18	53	30
8	35	12	35	13

APPENDIX E

Empirical Bayes Survival Analysis

Appendix E

Empirical Bayes Survival Analysis

The approach to the analysis of survival data that was used in this study is similar to that presented in Holford (1980) and Laird and Olivier (1981). The data are organized in a contingency table format with one factor representing time and the other factors representing the classification variables included in the analysis. Each cell contains information on the survival experience of a particular group of individuals over an interval of time. In this study, for example, the first series of survival analyses focused on the patterns of attainment for students in K departments within each of M schools.

A piecewise exponential distribution of survival times was assumed within each department at each school. This implies, for each of the KM departments, a constant hazard, θ_{ikm} (i = 1, 2...I; k = 1, 2,...K; m = 1, 2,...M), within each of I one-year time intervals. Let d_{ikmj} be an indicator variable such that $d_{ikmj} = 1$ if person j in department k at school m experiences the event (candidacy or graduation) in the i^{th} interval; otherwise, $d_{ikmj} = 0$. Let t_{ikmj} be the amount of time person j in department k at school m spends in the i^{th} interval. Let $d_{ikm} = \sum_{j=1}^{J} d_{ikmj}$ be the number of events that occur in interval i for department k at school m and let $t_{ikm} = \sum_{j=1}^{J} t_{ikmj}$ be the total exposure time in interval i for members of

department k at school m. We want to estimate the IKM parameters θ_{ikm}. The likelihood for this model is

$$\underset{\sim}{L}(\theta) = \prod_{m=1}^{M} \prod_{k=1}^{K} \prod_{j=1}^{J} \prod_{i=1}^{I} \theta_{ikm}^{d_{ikmj}} \exp(-\theta_{ikm} t_{ikmj})$$

$$= \prod_{m=1}^{M} \prod_{k=1}^{K} \prod_{i=1}^{I} \theta_{ikm}^{d_{ikm}} \exp(-\theta_{ikm} t_{ikm})$$

As demonstrated by Laird and Olivier (1981, p. 235) in the case of a simpler model, the likelihood obtained by assuming separate piecewise exponential distributions within the KM departments is proportional to the likelihood that would be obtained under the assumption that each d_{ikm} is an independent Poisson variate, conditional on t_{ikm}, with $E(d_{ikm}|t_{ikm}) = t_{ikm}\theta_{ikm}$. Because the likelihood kernels are the same, these two models can be used interchangeably for making likelihood-based inferences about the parameters θ_{ikm}. The maximum likelihood estimate of θ_{ikm} is simply the occurrence rate for department k at school k in interval i, d_{ikm}/t_{ikm}.

As a first step in the analysis, a "life table" was constructed for each department at each school, consisting of the values of d_{ikm} and n_{ikm}, where n_{ikm} is the number of students in department k at school m who had not yet experienced the event of interest as of the beginning of the i^{th} interval. A conventional life table formula was used to approximate the total exposure time for department k at school m in interval i:

$$\tilde{t}_{ikm} = n_{ikm} - \frac{d_{ikm} + c_{ikm}}{2} .$$

Here, c_{ikm} is the number of students in department k at school m who were censored during the i^{th} interval (see Laird & Oliver, p. 236). The assumption

underlying this approximation is that all events and censoring occur at the interval midpoint.

A problem with ratios of occurrence to exposure, like $d_{ikm}/\tilde{t}_{ikm}$ or d_{ikm}/t_{ikm}, is that they tend to be unstable when sample sizes are small. Empirical Bayes methodology provides a model-based approach to obtaining smoothed estimates of hazard and survival functions that are more stable than these conventional estimates and perform well on cross-validation (see Braun & Zwick, 1989). These more stable estimates for each of the KM departments can be obtained by incorporating information from other departments, which can be achieved quite naturally in the Bayesian framework by assuming a prior distribution for the θ_{ikm}. To remain in the Poisson framework, the next step would be to assume a prior distribution conjugate to the Poisson for the θ_{ikm}. The present approach, however, involves transforming the Poisson variates to normal variates and then applying empirical Bayes methods that have already been developed for the normal case. Let

$$X_{ikm} = [d_{ikm}/\tilde{t}_{ikm}]^{\frac{1}{2}}$$

Then, if the Poisson assumption holds, we have approximately

$$\underline{X}_{km} \sim N(\underline{\mu}_{km}, \underline{S}_{km})$$

where $\underline{X}_{km} = (X_{1km}, X_{2km}, \ldots X_{Ikm})$, $\underline{\mu}_{km} = (\theta^{\frac{1}{2}}_{1km}, \theta^{\frac{1}{2}}_{2km}, \ldots \theta^{\frac{1}{2}}_{Ikm})$ and $\underline{S}_{km}$ is a diagonal matrix with the i^{th} diagonal element equal to $(4\tilde{t}_{ikm})^{-1}$. The second level of the model assumes that

$$\underline{\mu}_{km} = \underline{G}\underline{Z}_{km} + \underline{D}$$

where $\underline{\mu}_{km}$ is the column vector of length I of transformed hazards for department k at school m. $\underline{Z}_{km}$ is a vector of length K + M - 1 that contains zeroes and ones indicating the department and school for the group under consideration and includes a one corresponding to the intercept of the model.

$\underline{G}$ is an I x (K + M - 1) matrix of coefficients to be estimated, and $\underline{D}$ is a vector of error components distributed as $N(\underline{0}, \underline{\Sigma}^*)$, where $\underline{\Sigma}^*$ is assumed to be unknown and must be estimated from the data.

This model is a special case of the general regression model described in Braun, Jones, Rubin, and Thayer (1983). Braun et al. show how the EM algorithm (Dempster, Laird, & Rubin, 1977) can be used to obtain maximum likelihood estimates of $\underline{G}$ and $\underline{\Sigma}^*$ as well as the posterior distributions of the $(\underline{\mu}_{km})$ given these estimates and the data. The means of these posterior distributions provide estimates of the $(\underline{\mu}_{km})$. Squaring these estimates in turn yields estimates of the $(\underline{\theta}_{km})$. (The estimation procedure for the empirical Bayes survival analysis model differs slightly from the general regression model in that the values of $Var(X_{ikm})$ are known in the survival analysis case and need not be reestimated in the M step of the EM algorithm.)

For the piecewise exponential survival model with intervals of length Δ_i, the probability of surviving through interval i_o for an individual in department k at school m is estimated by

$$\hat{S}_{km}(i_o) = \exp(-\sum_{i \le i_o} \hat{\theta}_{ikm}\Delta_i)$$

This expression is equal to $\exp(-\sum_{i \le i_o} \hat{\theta}_{ikm})$ if $\Delta_i = 1$, $i = 1, 2 \ldots I$, as in the present case. The empirical Bayes survival curves are obtained by setting $\hat{\theta}_{ikm}$ equal to the empirical Bayes estimates of the hazards.

The 11-department analyses in this study (see Figures 1 - 18) used a model identical to the one described above. That is, data from the 11 departments in the three schools were analyzed together, using department and school as factors in the analysis. The analyses of ethnic and gender groups

(see Figures 19-42) were conducted separately within each school, with department cluster and either ethnic or gender group as factors.

REFERENCES

American Council on Education. (1987). *Minorities in higher education*. Washington, DC: Author.

American Council on Education/Education Commission of the States. (1988). *One third of a nation: A report of the commission on minority participation in education and American life*. Washington, DC: American Council on Education.

Astin, A. W. (1982). *Minorities in American higher education*. San Francisco: Jossey-Bass.

Berg, H. M., & Ferber, M. A. (1983). Men and women graduate students: Who succeeds and why? *Journal of Higher Education, 54*, 629-648.

Blackwell, J. E. (1987). *Mainstreaming outsiders: The production of black professionals* (2nd ed.). Dix Hills, NY: General Hall.

Brademas, J. (1984). Graduate education: Signs of trouble. *Science, 223*, 881.

Braun, H. I. (1985). *Bayesian analysis of families of survival curves*. Unpublished paper, Educational Testing Service.

Braun, H. I. (1988). Empirical Bayes methods: A tool for exploratory analysis. In R. D. Bock (Ed.), *Proceedings of a conference on multilevel analysis of educational data*. New York: Academic Press. [Also available as Program Statistics Research Technical Report 88-82, Educational Testing Service, Princeton, NJ.]

Braun, H. I., Jones, D. H., Rubin, D. B., & Thayer, D. T. (1983). Empirical Bayes estimation of coefficients in the general linear model from data of deficient rank. *Psychometrika, 48*, 171-181.

Braun, H. I., & Zwick, R. (1989). *Bayesian analysis of families of survival curves: Applications to the analysis of degree attainment*. Unpublished manuscript, Educational Testing Service.

Brown, S. V. (1987). *Minorities in the graduate education pipeline*. (A research report of the Minority Graduate Education project, jointly sponsored by the GREB and ETS.) Princeton, NJ: Educational Testing Service.

Chamberlain, M. K. (Ed.) (1988). *Women in academe: Progress and prospects*. New York: Russell Sage Foundation.

Dawes, R. M. (1975). Graduate admissions variables and future success. *Science*, *187*, 721-723.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, B, *39*, 1-38.

Educational Testing Service. (1987). *1987-88 GRE guide to the use of the Graduate Records Examinations program*. Princeton, NJ: Author.

Feldman, S. D. (1974). *Escape from the doll's house: Women in graduate and professional school education*. New York: McGraw-Hill.

Girves, J. E., & Wemmerus, V. (1988). Developing models of graduate student progress. *Journal of Higher Education*, *59*, 163-189.

Hanushek, E. A., & Jackson, J. E. (1977). *Statistical methods for social scientists*. New York: Academic Press.

Hartnett, R. T. (1987). Has there been a graduate student "brain drain" in the arts and sciences? *Journal of Higher Education*, *58*, 562-585.

Holford, T. R. (1980). The analysis of rates and survivorship using loglinear models. *Biometrics*, *36*, 299-306.

Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Application to psychological measurement*. Homewood, IL: Dow Jones-Irwin.

Kalbfleisch, J. D., & Prentice, R. L. (1980). *The statistical analysis of failure time data*. New York: Wiley.

Laird, N. M., & Olivier, D. (1981). Covariance analyses of censored survival data using log-linear analysis techniques. *Journal of the American Statistical Association*, *76*, 231-240.

Mathematical Methods in Educational Research, School of Education, Stanford University. (1983). *A survival analysis of "Time-to-Ph.D." data for selected departments of Stanford University*. Report to Gerald R. Lieberman, Dean of Graduate Studies, Stanford University.

Mooney, C. J. (1989). Affirmative-action goals, coupled with tiny number of minority Ph.D.'s, set off faculty recruiting frenzy. *The Chronicle of Higher Education*, *35*, 1, 10-11.

Murnane, R. J., Singer, J. D., & Willett, J. B. (1988). The career paths of teachers: Implications for teacher supply and methodological lessons for research. *Educational Researcher*, *17*, 22-30.

National Board on Graduate Education. (1976). *Minority group participation in graduate education*. Washington, DC: National Academy of Sciences.

National Research Council. (1986). Summary report 1984: *Doctorate recipients from United States universities*. Washington, DC: National Academic Press.

Nettles, M. (1987). *Financial aid and minority participation in graduate education*. (A research report of the Minority Graduate Education [MGE] Project, jointly sponsored by the Graduate Record Examinations Board and Educational Testing Service.) Princeton, NJ: Educational Testing Service.

Patterson, M., & Sells, L. (1973). Women dropouts from higher education. In A. S. Rossi & A. Calderwood (Eds.), *Academic women on the move* (pp. 79-91). New York: Russell Sage Foundation.

Pruitt, A. S., & Isaac, P. D. (1985). Discrimination in recruitment, admission, and retention of minority graduate students. *Journal of Negro Education*, *54*, 526-536.

Roby, P. (1973). Institutional barriers to women students in higher education. In A. S. Rossi & A. Calderwood (Eds.), *Academic women on the move* (pp. 37-56). New York: Russell Sage Foundation.

Rubin, D. B. (1980). Using empirical Bayes techniques in the law school validity studies. *Journal of the American Statistical Association*, *75*, 801-816.

Schwartz, P., & Lever, J. (1973). Women in the male world of higher education. In A. S. Rossi & A. Calderwood (Eds.), *Academic women on the move* (pp. 57-77). New York: Russell Sage Foundation.

Thomas, G. E. (1987). Black students in U.S. graduate and professional schools in the 1980's: A national and institutional assessment. *Harvard Educational Review*, *57*, 261- 282.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, *45*, 89-125.

Trent, W. T., & Copeland, E. J. (1987). *Effectiveness of state financial aid in the production of Black doctoral recipients*. Atlanta: Southern Education Foundation.

Widnall, S. E. (1988). AAAS presidential lecture: Voices from the pipeline. *Science*, *241*, 1740-1745.

Willingham, W. W. (1974). Predicting success in graduate education. *Science*, *183*, 273-278.

Willingham, W. W. (1985). *Success in college: The role of personal qualities and academic ability*. New York: College Entrance Examination Board.

Zwick, R. (1990). *The validity of the GMAT for the prediction of success in doctoral study in business and management*. (ETS Research Report 90-24.) Princeton, NJ: Educational Testing Service.

Zwick, R., & Braun, H. I. (1988). *Methods for analyzing the attainment of graduate school milestones: A case study*. (GRE Board Professional Report 86-3P, ETS Research Report 88-30.) Princeton, NJ: Educational Testing Service.

54020-01161 • S61M.7 • 209137

BEST COPY AVAILABLE